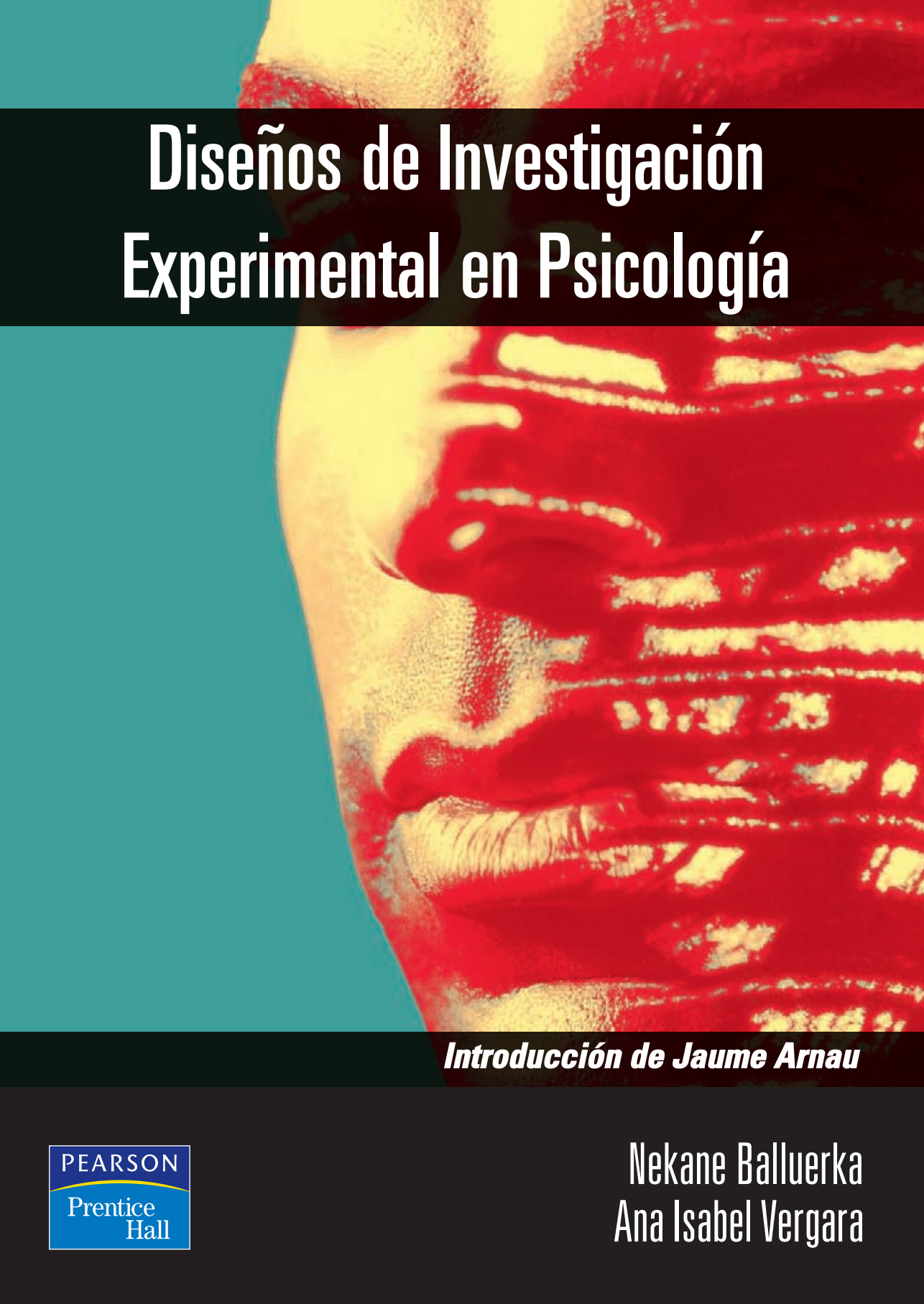




Diseños de Investigación Experimental en Psicología

Introducción de Jaume Arnau



Nekane Balluerka
Ana Isabel Vergara

**DISEÑOS DE INVESTIGACIÓN
EXPERIMENTAL EN PSICOLOGÍA
MODELOS Y ANÁLISIS DE DATOS MEDIANTE EL SPSS 10.0**

DISEÑOS DE INVESTIGACIÓN EXPERIMENTAL EN PSICOLOGÍA

MODELOS Y ANÁLISIS DE DATOS MEDIANTE EL SPSS 10.0

Nekane Balluerka Lasa

Profesora Titular de Psicología
Universidad del País Vasco

Ana Isabel Vergara Iraeta

Profesora de Psicología
Universidad del País Vasco

Introducción:

Jaume Arnau y Gras

Catedrático de Psicología
Universidad de Barcelona



Madrid • México • Santafé de Bogotá • Buenos Aires • Caracas • Lima • Montevideo • San Juan
San José • Santiago • São Paulo • White Plains

Nekane Balluerka Lasa, Ana Isabel Vergara Iraeta

Diseños de Investigación Experimental en Psicología

PEARSON EDUCACIÓN, Madrid, 2002

ISBN: 978-84-205-3447-3

Materia: Psicología 159.9

Formato 170 × 240

Páginas: 432

No está permitida la reproducción total o parcial de esta obra ni su tratamiento o transmisión por cualquier medio o método, sin autorización escrita de la Editorial.

DERECHOS RESERVADOS

© 2002 PEARSON EDUCACIÓN, S. A.

Ribera del Loira, 28

28042 MADRID

Nekane Balluerka Lasa, Ana Isabel Vergara Iraeta

Diseños de investigación experimental en Psicología

ISBN: 978-84-205-3447-3

Depósito legal: M.

PRENTICE HALL es un sello editorial autorizado de PEARSON EDUCACIÓN, S. A.

Editor: Juan Luis Posadas

Equipo de producción:

Director: José Antonio Clares

Técnico: Isabel Muñoz

Diseño de cubierta: Mario Guindel, Lía Sáenz y Begoña Pérez

Composición: COPIBOOK, S. L.

Impreso por:

IMPRESO EN ESPAÑA - PRINTED IN SPAIN



*Gorka, Amaia eta Eiderentzat
Bittor eta Gontzalen baimenarekin.*

*A Gorka, Amaia y Eider
con permiso de Bittor y Gontzal.*

AGRADECIMIENTOS

Queremos hacer constar nuestro agradecimiento a todas aquellas personas que, de uno u otro modo, han contribuido a la elaboración de este trabajo. En especial, al profesor Jaume Arnau quien, además de acceder a realizar la introducción de este texto, nos ha aportado sus conocimientos, así como su constante apoyo y disponibilidad. Al profesor Manuel Ato, quien ha llevado a cabo una revisión exhaustiva del libro y nos ha hecho sugerencias realmente interesantes para mejorar la calidad del mismo. A nuestro compañero del área de Metodología, Xabier Isasi, por el tiempo invertido en enriquecedoras discusiones y por su continuo aliento.

Queremos también agradecer la colaboración de Estefanía Ocariz y Beatriz Rodríguez en la transcripción de diversas partes del texto. A su vez, sería injusto olvidar a nuestros alumnos quienes, en gran medida, han motivado la elaboración de este trabajo. También queremos mostrar nuestro agradecimiento a la editorial Prentice Hall y, en especial, a quienes en ella han confiado y apostado por nuestro trabajo.

Finalmente, queremos dar las gracias a todos los que nos han apoyado, tanto desde el ámbito universitario como fuera de éste, y que no han sido nombrados en estas líneas, en especial a Bittor y a Gontzal quienes han cedido el protagonismo que, por méritos propios, les correspondía en la dedicatoria de este libro.

TABLA DE CONTENIDOS

INTRODUCCIÓN	xv
CAPÍTULO 1: CONCEPTO DE DISEÑO	1
CAPÍTULO 2: PRINCIPALES ALTERNATIVAS METODOLÓGICAS Y DISEÑOS EN PSICOLOGÍA: UNA PERSPECTIVA GENERAL	5
2.1. Principales estrategias para la recogida de los datos	5
2.2. Diseños experimentales	8
2.3. Diseños cuasi-experimentales	9
2.4. Diseños no-experimentales	10
CAPÍTULO 3: ASPECTOS ESENCIALES DE LA METODOLOGÍA EXPERIMENTAL ..	13
3.1. Definición y objetivos de la experimentación	13
3.1.1. Paradigma experimental contra paradigma asociativo	13
3.1.2. Concepto de experimento	14
3.1.3. Objetivos del experimento	16
3.2. Criterios que garantizan la calidad metodológica del experimento: el principio MAX-MIN-CON	18
CAPÍTULO 4: PRINCIPALES CRITERIOS PARA LA CLASIFICACIÓN DE LOS DISEÑOS EXPERIMENTALES CLÁSICOS O FISHERIANOS	21
CAPÍTULO 5: APROXIMACIÓN GENERAL A LA TÉCNICA DE ANÁLISIS DE DATOS MÁS UTILIZADA EN EL ÁMBITO DEL DISEÑO EXPERIMENTAL CLÁSICO: EL ANÁLISIS DE LA VARIANZA	27
5.1. Análisis univariado de la varianza: ANOVA	27
5.1.1. Sumas de cuadrados	30
5.1.2. Grados de libertad	32
5.1.3. Varianzas o medias cuadráticas	32
5.1.4. Razón F	33

5.2.	Análisis multivariado de la varianza: MANOVA	34
5.2.1.	Relaciones entre el ANOVA y el MANOVA	34
5.2.2.	Prueba de la hipótesis de nulidad en los diseños multivariados	35
	• Criterio de la razón de verosimilitud	38
	• Criterio de la raíz más grande	39
	• Criterio de la traza	39
CAPÍTULO 6: DISEÑOS UNIFACTORIALES ALEATORIOS		43
6.1.	Diseño de dos grupos aleatorios	43
6.1.1.	Características generales del diseño de dos grupos aleatorios	43
6.1.2.	El análisis de datos en el diseño de dos grupos aleatorios	44
	6.1.2.1. Modelo general de análisis	44
	6.1.2.2. Ejemplo práctico	47
	• Prueba de Hartley	47
	• t de Student	48
	• Análisis unifactorial de la varianza	48
	• Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	54
6.2.	Diseño multigrupos aleatorios	57
6.2.1.	Características generales del diseño multigrupos aleatorios	57
6.2.2.	El análisis de datos en el diseño multigrupos aleatorios	58
	6.2.2.1. Posibilidades analíticas para el diseño multigrupos aleatorios	58
	6.2.2.2. El análisis unifactorial de la varianza	58
	6.2.2.3. Ejemplo práctico del análisis unifactorial de la varianza	59
	• Prueba de Hartley	60
	• Análisis unifactorial de la varianza	60
	6.2.2.4. Comparaciones múltiples entre medias	67
	A. Elección del procedimiento para la realización de las comparaciones múltiples	68
	B. Principales procedimientos de comparaciones múltiples: ejemplos prácticos	77
	• Corrección de Bonferroni	77
	• Procedimiento de Dunnett	81
	• Procedimiento DHS de Tukey	82
	• Procedimiento de Scheffé	84
	6.2.2.5. Análisis de datos mediante el paquete estadístico SPSS 10.0	84
CAPÍTULO 7: DISEÑOS FACTORIALES ALEATORIOS		89
7.1.	Principales criterios para la clasificación de los diseños factoriales	89
7.2.	Ventajas del diseño factorial frente al diseño unifactorial	90
7.3.	El análisis de datos en los diseños factoriales aleatorios	92
	7.3.1. Posibilidades analíticas para los diseños factoriales aleatorios	92
	7.3.2. Análisis factorial de la varianza para el diseño factorial $A \times B$	93
	7.3.2.1. Modelo general de análisis	93
	7.3.2.2. Ejemplo práctico	94
	7.3.2.3. Desarrollo del análisis factorial de la varianza	96
	7.3.2.4. Análisis de datos mediante el paquete estadístico SPSS 10.0	106
	7.3.3. Análisis factorial de la varianza para el diseño factorial $A \times B \times C$	113
	7.3.3.1. Modelo general de análisis	113
	7.3.3.2. Ejemplo práctico	114
	7.3.3.3. Desarrollo del análisis factorial de la varianza	116
	7.3.3.4. Análisis de datos mediante el paquete estadístico SPSS 10.0	137

7.3.4.	El estudio de los efectos de interacción: análisis de los efectos simples	142
7.3.4.1.	Ejemplo práctico	143
7.3.5.	Comparaciones múltiples entre medias para diseños factoriales con efecto de interacción	151
7.3.5.1.	Corrección de Bonferroni	152
7.3.5.2.	Procedimiento DHS de Tukey	153
7.3.5.3.	Prueba de Scheffé	155
7.3.5.4.	Análisis de datos mediante el paquete estadístico SPSS 10.0	156

CAPÍTULO 8: DISEÑOS EXPERIMENTALES QUE REDUCEN LA VARIANZA DE ERROR		157
8.1.	Diseños con un factor de bloqueo: diseños de bloques aleatorios	158
8.1.1.	Características generales del diseño de bloques aleatorios	158
8.1.2.	Análisis factorial de la varianza para el diseño de bloques aleatorios	159
8.1.2.1.	Análisis factorial de la varianza para el diseño de un solo sujeto por casilla o por combinación entre tratamiento y bloque	159
	• Modelo general de análisis	159
	• Ejemplo práctico	160
8.1.2.2.	Análisis factorial de la varianza para el diseño con más de un sujeto por casilla o por combinación entre tratamiento y bloque	164
	• Modelo general de análisis	164
	• Ejemplo práctico	165
	• Desarrollo del análisis factorial de la varianza	167
	• Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	177
8.2.	Diseños con dos factores de bloqueo: diseños de cuadrado latino	180
8.2.1.	Características generales del diseño de cuadrado latino	180
8.2.2.	Análisis factorial de la varianza para el diseño de cuadrado latino	181
8.2.2.1.	Modelo general de análisis	181
8.2.2.2.	Ejemplo práctico	183
8.2.2.3.	El problema de la reducción de los grados de libertad del término de error	186
8.2.2.4.	Análisis de datos mediante el paquete estadístico SPSS 10.0	187
8.3.	Diseños jerárquicos	190
8.3.1.	Características generales del diseño jerárquico	190
8.3.2.	Análisis factorial de la varianza para el diseño jerárquico	191
8.3.2.1.	Modelo general de análisis	191
8.3.2.2.	Consideraciones importantes con respecto a la selección del término de error para contrastar el efecto de cada uno de los factores incluidos en el diseño	192
	• Diseño jerárquico de dos factores con un factor de anidamiento y un factor anidado	194
	• Diseño jerárquico de tres factores con un factor de anidamiento, un factor anidado y un factor que mantiene una relación de cruce con el factor anidado	195
	• Diseño jerárquico de tres factores con un factor de anidamiento y dos factores anidados	196
8.3.2.3.	Ejemplo práctico	197
8.3.2.4.	Análisis de datos mediante el paquete estadístico SPSS 10.0	205
8.4.	Diseños con covariables	208
8.4.1.	Características generales del diseño con covariables	208
8.4.2.	Diseño totalmente aleatorio, diseño de bloques aleatorios y diseño con covariables: breve análisis comparativo	210

8.4.3.	El análisis de datos en los diseños con covariables: análisis de la covarianza (ANCOVA)	210
8.4.3.1.	Supuestos básicos del análisis de la covarianza	210
8.4.3.2.	Modelo general de análisis	213
8.4.3.3.	Ejemplo práctico	214
8.4.3.4.	Análisis de datos mediante el paquete estadístico SPSS 10.0	234
CAPÍTULO 9: DISEÑOS EXPERIMENTALES DE MEDIDAS REPETIDAS		241
9.1.	Diseños simples y factoriales de medidas totalmente repetidas	241
9.1.1.	Características generales del diseño de medidas repetidas	241
9.1.2.	El análisis de datos en los diseños de medidas totalmente repetidas	243
9.1.2.1.	Supuestos básicos para el análisis y alternativas ante su incumplimiento	243
9.1.2.2.	Análisis de la varianza mixto para el diseño intrasujeto simple (diseño de tratamientos $\times$ sujetos)	246
	• Modelo general de análisis	246
	• Ejemplo práctico	247
	• Desarrollo del análisis de la varianza mixto para el diseño intrasujeto simple	249
	• Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	256
9.1.2.3.	Análisis de la varianza mixto para el diseño factorial intrasujeto de dos factores (diseño de tratamientos $\times$ tratamientos $\times$ sujetos)	259
	• Modelo general de análisis	259
	• Ejemplo práctico	260
	• Desarrollo del análisis de la varianza mixto para el diseño intrasujeto de dos factores	262
	• Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	275
9.1.2.4.	Comparaciones múltiples entre medias	279
9.2.	Diseños de medidas parcialmente repetidas: diseño factorial mixto y diseño <i>split-plot</i> ..	282
9.2.1.	Características generales del diseño factorial mixto y del diseño <i>split-plot</i> ..	282
9.2.2.	El análisis de la varianza mixto para los diseños de medidas parcialmente repetidas	283
9.2.2.1.	Supuestos básicos para el análisis de los diseños de medidas parcialmente repetidas	283
9.2.2.2.	Modelo general de análisis	283
9.2.2.3.	Ejemplo práctico: Diseño <i>split-plot</i>	285
9.2.2.4.	Desarrollo del análisis de la varianza mixto para el diseño de medidas parcialmente repetidas	286
9.2.2.5.	Comparaciones múltiples entre medias	296
9.2.2.6.	Análisis de datos mediante el paquete estadístico SPSS 10.0	300
9.3.	Diseño <i>cross-over</i> o conmutativo y diseño de cuadrado latino intrasujeto	303
9.3.1.	Diseño <i>cross-over</i> o conmutativo	303
9.3.1.1.	Características generales del diseño <i>cross-over</i>	303
9.3.1.2.	El análisis de la varianza para el diseño <i>cross-over</i>	303
	• Modelo general de análisis	303
	• Ejemplo práctico	304
	• Desarrollo del análisis de la varianza para el diseño <i>cross-over</i> 2×2	305
	• Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	310
9.3.2.	Diseño de cuadrado latino intrasujeto	313
9.3.2.1.	Características generales del diseño de cuadrado latino intrasujeto	313

9.3.2.2.	El análisis de la varianza para el diseño de cuadrado latino intrasujeto	314
•	Modelo general de análisis	314
•	Ejemplo práctico	315
•	Análisis de datos mediante el paquete estadístico SPSS 10.0 ...	317
CAPÍTULO 10:	OTRAS MODALIDADES DE DISEÑO	321
10.1.	Diseño experimental multivariado	321
10.1.1.	Características generales del diseño experimental multivariado	321
10.1.2.	El análisis de datos en el diseño multivariado: análisis multivariado de la varianza (MANOVA)	322
10.1.2.1.	Consideraciones generales acerca del modelo multivariante ...	322
10.1.2.2.	Ejemplo práctico	323
10.1.2.3.	Análisis de datos mediante el paquete estadístico SPSS 10.0 ..	329
10.2.	Diseño de bloques incompletos	332
10.2.1.	Características generales del diseño de bloques incompletos	332
10.2.2.	La técnica de confusión	333
10.2.2.1.	Técnica de confusión completa	333
10.2.2.2.	Técnica de confusión parcial	336
10.2.3.	El análisis de la varianza para el diseño de bloques incompletos	338
10.2.3.1.	Ejemplo práctico	338
10.2.3.2.	Desarrollo del análisis de la varianza para el diseño de bloques incompletos 2×2 con la interacción $A \times B$ totalmente confundida	340
10.2.3.3.	Análisis de datos mediante el paquete estadístico SPSS 10.0 ..	341
10.3.	Diseño factorial fraccionado	346
10.3.1.	Características generales del diseño factorial fraccionado	346
10.3.2.	La técnica de replicación fraccionada	347
10.3.3.	El análisis de la varianza para el diseño factorial fraccionado	350
10.3.3.1.	Ejemplo práctico	350
10.3.3.2.	Desarrollo del análisis de la varianza para el diseño factorial fraccionado	350
10.3.3.3.	Análisis de datos mediante el paquete estadístico SPSS 10.0 ..	352
ANEXO A:	TABLAS ESTADÍSTICAS	357
ANEXO B:	INTRODUCCIÓN DE DATOS EN EL EDITOR DEL SPSS 10.0	379
REFERENCIAS BIBLIOGRÁFICAS		391
ÍNDICE DE TÉRMINOS		399

INTRODUCCIÓN

Desde que la psicología devino en ciencia objetiva o en disciplina de base empírica, se ha asistido a un constante desarrollo y evolución de los procedimientos de estudio y análisis. Ello ha propiciado el avance de la psicología científica que, sin duda, se ha beneficiado del progreso en las técnicas de diseño y análisis de datos. Nótese, por otra parte, que el objetivo fundamental de la investigación psicológica radica en ampliar el conocimiento sobre el comportamiento humano, entendido en un sentido global.

La investigación suele ser, por lo general, concebida en función de cuatro objetivos fundamentales, jalonados en estadíos sucesivos: el primero consiste en describir, el segundo en comprender, el tercero en predecir, y el último en controlar (Clark-Carter, 1998). Cada uno de estos cuatro objetivos se alcanza mediante una determinada metodología. Así, mediante la metodología observacional se describe la realidad de los hechos o de las dimensiones de variación del objeto de estudio que, en psicología, es la conducta o el comportamiento, en su sentido más amplio. La descripción se amplía a partir del conocimiento y de la medición de estas dimensiones o variables, y de sus hipotéticas relaciones, con el propósito de predecir. Esta es la función básica de la metodología de encuesta. Mediante la metodología experimental se verifica el efecto causal de los tratamientos sobre la conducta, así como el control que éstos pueden ejercer sobre ella. En este contexto, lo que controla, causa, y lo que causa, explica. De ahí, que mediante la metodología experimental es posible conocer la función explicativa de las variables. Por último, la metodología cuasi-experimental, si bien participa del propósito de la metodología experimental, permite evaluar el resultado de la intervención en sujetos o sistemas sociales amplios. Ha de quedar claro que las distintas metodologías constituyen, dentro del modelo general de investigación, el puente que une lo conceptual con lo empírico y aportan el fundamento objetivo a los modelos teóricos.

Centrándonos en la materia fundamental del texto, cabría preguntarse qué es lo que se entiende por metodología experimental. La metodología experimental, como cualquier otro enfoque de estudio, forma parte de un proceso o modelo general, conocido como método científico, que tiene por objeto la consecución de conocimientos con significación empírica. La investigación psicológica, sin el uso del método científico, se reduciría a meras especu-

laciones de carácter filosófico o literario. Sin método no hay ciencia y sin ciencia no hay explicación causal o teoría explicativa sobre la realidad. De ahí la necesidad de un método que sea común a todas las ciencias positivo-naturales. Se trata, por tanto, de una estrategia particular de investigación que reproduce el proceso o modelo general en una situación concreta.

Desde una perspectiva metateórica, las distintas estrategias de investigación se inscriben dentro de un conjunto de marcos de actuación o *marcos metodológicos* que definen los objetivos, su consecución y los procedimientos de obtención de datos. Tales marcos se conocen como *paradigmas*. Los paradigmas asumen, entre otras cosas, un conjunto de postulados metateóricos y metodológicos que dictan las reglas, tanto para la construcción de los esquemas explicativos como para los procedimientos de investigación. Sin pretender remontarnos a los enunciados de carácter metateórico en el sentido kuhniano, el término paradigma puede tomarse como sinónimo de *sistema inspirador de metodologías de trabajo*. En función de esta caracterización de paradigma, en la ciencia psicológica están presentes dos paradigmas o tradiciones: el paradigma experimental y el paradigma asociativo. Cada paradigma se caracteriza por la formulación de una clase específica de hipótesis, por el grado de intervención del investigador en la situación estudiada, por los sistemas de recogida de datos, y por los procedimientos de verificación de las hipótesis (Kuhn, 1962, 1970; Madsen, 1978, 1980).

A modo de resumen, cabe concluir que las metodologías que configuran el panorama actual de la investigación psicológica no surgen de un vacío conceptual, ni aparecen de forma espontánea dentro del ámbito de los procedimientos de estudio. Son estrategias que se derivan de una serie de postulados metateóricos que, por otra parte, trascienden el ámbito meramente fenoménico y experiencial. Es interesante tener en cuenta esta perspectiva, ya que los procedimientos de estudio constituyen la respuesta metodológica a los problemas que están más allá de la experiencia directa de los hechos y de los datos (Arnau y Balluerka, 1998).

MODELO GENERAL DE INVESTIGACIÓN

El modelo general de la investigación, dentro del contexto de las ciencias psicológicas y sociales, está estructurado en tres niveles (Arnau, 1990, 1995), formalmente jerarquizados: nivel de expectativa teórica, nivel de diseño y nivel de análisis de datos, tal y como se muestra en la Figura 1.

Según el modelo general de la Figura 1, el primer nivel define el carácter conceptual del modelo. En este nivel se delimitan los ámbitos de la experiencia observada, se formulan problemas o cuestiones acerca de lo observado, se plantean posibles explicaciones teóricas en términos de hipótesis o conjeturas, y se derivan las consecuencias empíricamente contrastables a partir de las hipótesis. Estas consecuencias son, de hecho, las hipótesis de trabajo sobre las que se planifica el diseño, en función de una determinada estrategia de recogida de datos. Así, en el nivel de expectativas teóricas, el modelo general se centra fundamentalmente en las hipótesis o expectativas acerca de la forma en la que se articulan las variables, así como en los posibles nexos existentes entre los fenómenos observados que definen el referente empírico de una ciencia.

En un segundo nivel, caracterizado por la operativización de la hipótesis, se definen las estrategias que se van a utilizar para obtener la información requerida y para resolver los problemas planteados en el nivel teórico. La tarea más importante en este segundo nivel, de carácter técnico y práctico, es la identificación de las variables que conforman la situación de investigación. Sobre cada una de estas variables, el investigador decide el grado de inter-

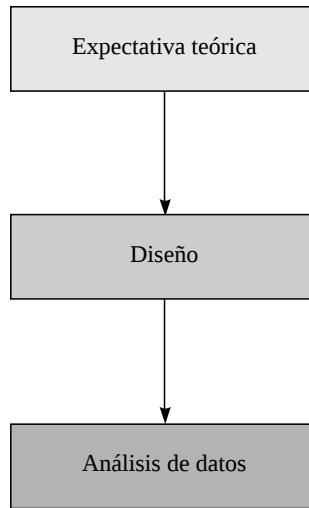


Figura 1 Modelo general de la investigación psicológica y social.

vención, es decir, si ha de manipularla, medirla o simplemente neutralizarla mediante las distintas técnicas de control. El objetivo de este segundo nivel radica en conseguir información sobre las distintas variables implicadas en el diseño de estudio o de trabajo.

En el nivel de análisis, el investigador recurre a los modelos estadísticos y a sus correspondientes pruebas de significación. Para ello se parte de la hipótesis de nulidad estadística con el propósito de rechazarla. En función del resultado obtenido en la prueba estadística, se interpretan los datos a la luz de los supuestos y de los modelos teóricos que han activado el proceso de investigación. Téngase en cuenta que el modelo de análisis depende de dos consideraciones fundamentales: la naturaleza de los datos (datos cuantitativos o categóricos) y la estructura del diseño (cantidad de grupos y variables independientes).

De todo lo expuesto hasta el momento, cabe concluir que el modelo general de investigación es un proceso de carácter secuencial, estructurado en diferentes niveles. Cada nivel se rige por una serie de presupuestos y de reglas referidas a la función que desempeña dentro del proceso en general. A su vez, las metodologías de trabajo son procedimientos concretos que resuelven los problemas de investigación y constituyen el vehículo de enlace entre el modelo general de investigación y una determinada situación de estudio. Si bien la ciencia psicológica cuenta con una gran variedad de metodologías, éstas se distinguen en función del grado de rigor y potencia inferencial, es decir, del grado en que se controlan las causas o variables que afectan a la variación de la variable dependiente u objeto de análisis.

METODOLOGÍA DE INVESTIGACIÓN EXPERIMENTAL Y CUASI-EXPERIMENTAL

El paradigma experimental, a nivel metateórico, parte de una serie de presupuestos que rigen las estrategias de diseño y los procedimientos de trabajo. Entre tales presupuestos cabe destacar la asunción de relaciones causales entre las variables, la asincronía temporal entre la variable causa y la variable efecto, la proporcionalidad entre la variación de la variable independiente y la variable dependiente, y la covariabilidad entre ambas. Tanto la estrategia

experimental como la cuasi-experimental participan de estos presupuestos, de modo que los distintos diseños derivados de ambos enfoques se ajustan fundamentalmente a tales condiciones. Nuestro interés va a centrarse, a lo largo de los párrafos siguientes, en destacar las diferencias existentes entre estas dos estrategias de investigación, lo que nos permitirá comprender más adecuadamente el carácter particular de los diseños tanto experimentales como cuasi-experimentales.

La metodología experimental, directamente asociada a las ciencias positivo-naturales, persigue el objetivo de estimar la magnitud del efecto de las variables de tratamiento o variables independientes (variables causa). Desde el punto de vista formal, esta metodología es la fiel expresión del paradigma experimental, ya que en ella se materializan sus principales presupuestos. La metodología experimental se fundamenta en tres grandes pilares o componentes básicos: aleatorización, control e hipótesis de causalidad (Gad, 1999).

Como se muestra en la Figura 2, la aleatorización, el control y el modelo de causalidad, en virtud de la manipulación directa de la variable causa, son las tres señas de identidad de la metodología experimental. La aleatorización sirve, en primer lugar, para homogeneizar los grupos y hacerlos equivalentes antes de aplicar el tratamiento. El control, en segundo lugar, garantiza la neutralización de cualquier fuente sistemática de variación susceptible de rivalizar con la variable causa. Un principio básico del control consiste en la asignación aleatoria de los sujetos a los distintos grupos experimentales, de ahí que la investigación experimental esté asociada al principio de la aleatorización de los sujetos (Chow y Liu, 1998). El modelo de causalidad permite, en tercer lugar, conocer de forma directa cuál ha sido la variable, previamente manipulada, que ha causado la variación en los datos u observaciones de los sujetos.

Más concretamente, la investigación experimental persigue los siguientes propósitos (Petersen, 1985):

- a) Proporcionar estimaciones de los efectos de los tratamientos o de las diferencias entre los efectos de los tratamientos.

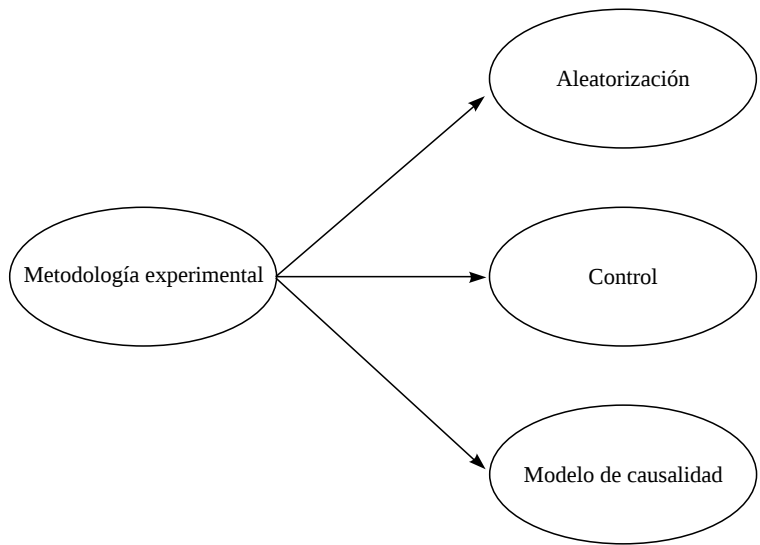


Figura 2 Componentes básicos de la metodología experimental.

- b) Aportar un sistema eficaz para confirmar o rechazar conjeturas sobre la posible respuesta al tratamiento.
- c) Evaluar la fiabilidad de las estimaciones y conjeturas planteadas.
- d) Estimar la variabilidad de las unidades o de los datos experimentales.
- e) Aumentar la precisión en la estimación, eliminando, de las comparaciones de interés, la variación debida a variables extrañas.
- f) Aportar un modelo sistemático y eficiente para llevar a cabo un experimento.

¿Cuál sería la alternativa a la investigación experimental, dentro del paradigma experimental? Hay situaciones, fuera del contexto de laboratorio, donde ni la aleatorización de las unidades o de los sujetos es posible, ni el control de las fuentes alternativas de explicación resulta efectivo. Estas situaciones son externas al laboratorio y no hay posibilidad de asignar los sujetos a las condiciones de tratamiento o a los grupos. Se trata, en este caso, de trabajar con grupos ya formados o naturales como, por ejemplo, las aulas de un centro escolar, o las plantas de un hospital. Bajo estas circunstancias particulares, la mejor alternativa a la metodología experimental es la investigación cuasi-experimental, de carácter fundamentalmente aplicado.

La metodología de investigación cuasi-experimental se utiliza para estudiar el posible efecto causal de las intervenciones o de los tratamientos en situaciones abiertas, es decir, fuera del contexto del laboratorio, donde el control es escaso y la aleatorización en la asignación de las unidades no resulta posible. Por lo general, los diseños cuasi-experimentales plantean cuestiones prácticas que tienen interés en distintos contextos de aplicación. Tales diseños fueron sistematizados por primera vez en los textos de Campbell y Stanley (1966), y de Cook y Campbell (1979). Cabe insistir, de nuevo, en que los diseños cuasi-experimentales parten de grupos que ya están formados o bien son grupos naturales y que, en muchas situaciones, se desconoce cuál es la población de origen. Por esta razón, en múltiples contextos, tales diseños se conocen también como diseños no aleatorizados, en oposición a los diseños experimentales o aleatorizados.

La Figura 3 muestra los componentes básicos que configuran la investigación cuasi-experimental. El primer componente, que imprime carácter a este procedimiento, hace referencia a la ausencia de aleatorización en la formación de los grupos. En esta metodología los grupos ya están formados, es decir, son grupos naturales o intactos. Ocurre, frecuentemente, que se desconoce incluso su procedencia. En todo caso, la ausencia de aleatorización en la formación de los grupos los convierte en no comparables o no equivalentes inicialmente, lo cual complica el modelo estadístico para la prueba de la hipótesis. En segundo lugar, dada la ausencia del azar en la asignación de las unidades, es posible que algunas variables extrañas de confundido escapen al control y se conviertan en fuente de sesgo de los grupos. Por último, como resultado de lo anterior, es posible que existan modelos explicativos que rivalicen con el modelo de la hipótesis causal. El desarrollo de esta nueva metodología de trabajo se ha visto potenciado por la necesidad de llevar a cabo estudios de carácter aplicado dentro de la psicología social y educativa. Especialmente en áreas donde el interés radica en evaluar el posible efecto de determinados programas sociales o innovaciones políticas, lo que ha dado lugar a la metodología de evaluación de programas. El principal problema asociado a la metodología aplicada o cuasi-experimental es el relativo a la inferencia de la causalidad. Ello se debe a que la amenaza más importante que se cierne sobre tales estudios es la posibilidad de que una tercera variable explique la relación observada entre el tratamiento y la variable de respuesta, o el hecho de que los datos sean asumibles por más de un modelo explicativo rival. De ahí la estrecha vinculación que existe entre asignación aleatoria de las unidades e inferencia de relaciones casuales.

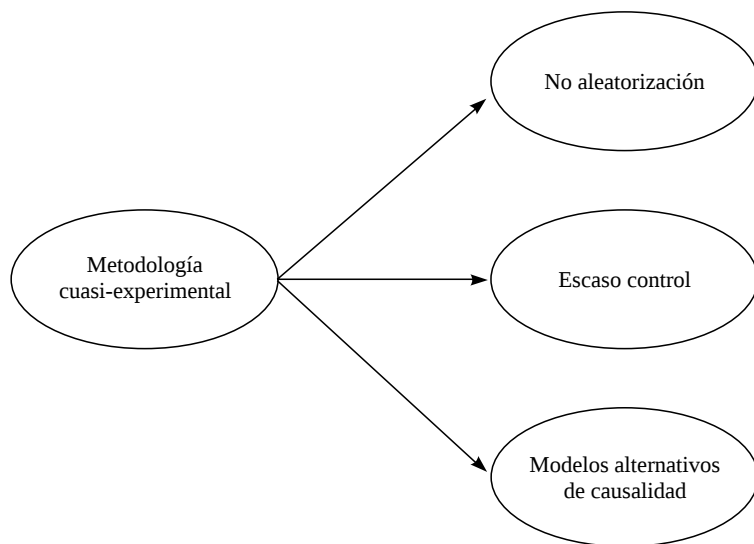


Figura 3 Componentes básicos de la metodología cuasi-experimental.

La investigación cuasi-experimental tiene como objetivos básicos los siguientes (Hedrick, Bickman y Rog, 1993; Pedhazur y Schmelkin, 1991; Ross y Grant, 1994):

- a) Estudiar el efecto de las variables de tratamiento o de las intervenciones en aquellas situaciones en las que los sujetos no han sido asignados aleatoriamente a los grupos.
- b) Evitar, en la medida de lo posible, el error de especificación, es decir, la omisión de variables correlacionadas con la variable de tratamiento.
- c) Identificar las variables relacionadas con la independiente y tenerlas en cuenta en el análisis, a fin de que las estimaciones de los efectos no resulten sesgadas.
- d) Corregir, mediante el modelo estadístico, el sesgo que presentan los grupos debido a su origen.
- e) Seleccionar el modelo estadístico más adecuado en función de la estructura del diseño, y obtener una inferencia válida.
- f) En situaciones longitudinales, estudiar los procesos de cambio y las posibles causas de dicho cambio.

Cabe insistir, de nuevo, en que el enfoque cuasi-experimental, a diferencia del enfoque experimental, no asume la equivalencia inicial entre los grupos, es decir, debido a la ausencia de aleatorización de las unidades, los grupos suelen ser diferentes antes de la aplicación de los tratamientos. Estas diferencias constituyen el principal obstáculo para poder inferir un resultado válido.

PERSPECTIVA GENERAL DE LOS DISEÑOS DE INVESTIGACIÓN EXPERIMENTALES Y CUASI-EXPERIMENTALES EN CIENCIAS DEL COMPORTAMIENTO

En esta introducción se ha seguido como hilo argumental la idea de que los paradigmas metodológicos definen y guían los sistemas de trabajo o las estrategias de recogida de los

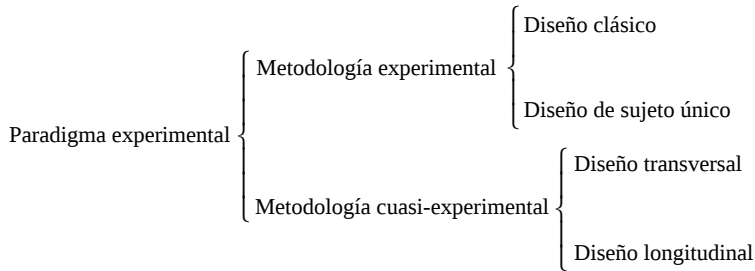


Figura 4 Clasificación de las metodologías y de los diseños del paradigma experimental.

datos, es decir, los diseños de investigación. Desde esta perspectiva, se describirán de forma sucinta las principales modalidades de diseño, tanto en su versión experimental como cuasi-experimental, en función de los criterios que articulan su categorización. Estos criterios varían según el diseño sea experimental o cuasi-experimental.

Considerando, en primer lugar, el diseño experimental o basado en el azar, cabe plantear como criterio de clasificación el hecho de que la muestra esté formada por un conjunto de sujetos o por un solo sujeto. En función de este criterio, el diseño experimental se clasifica en dos categorías: diseño experimental clásico y diseño conductual o de sujeto único. En el caso del diseño cuasi-experimental, el criterio de clasificación es el procedimiento utilizado para la comparación entre las observaciones. Así, las comparaciones pueden ser de carácter estático, como en el diseño denominado transversal, o de carácter dinámico, en cuyo caso el diseño es considerado longitudinal.

Como cabe observar en la Figura 4, las metodologías más importantes del paradigma experimental incluyen diferentes estrategias de recogida de datos o de diseño. Aunque no hay una correspondencia exacta entre las estrategias del enfoque experimental y las estrategias del enfoque cuasi-experimental, cabe destacar algunas semejanzas. Así, el diseño experimental clásico es análogo, en su estructura, al diseño cuasi-experimental transversal, ya que ambas estrategias resuelven la inferencia mediante la comparación entre grupos. Por esa razón tales diseños son también conocidos como diseños de grupos paralelos. En cuanto al diseño de sujeto único de carácter experimental y al diseño longitudinal de carácter cuasi-experimental, ambos se caracterizan por el hecho de repetir medidas en una o más unidades de observación. Por tal razón, la dimensión temporal es el atributo que mejor define a esa clase de diseños.

DOS TRADICIONES EN INVESTIGACIÓN EXPERIMENTAL

La investigación experimental, que constituye el centro de interés del presente manual, ha seguido a lo largo de la pasada década dos líneas de desarrollo conocidas como enfoque clásico y enfoque conductual o de sujeto único (figura 4). Esto ha configurado dos grandes tradiciones de diseño que, históricamente, han marcado el devenir de las ciencias psicológicas y sociales. El enfoque clásico parte de los trabajos de Fisher (1918, 1925, 1935), donde se introducen los conceptos de variancia y análisis de la variancia, así como el concepto de diseño experimental. La gran aportación de Fisher (1935), dejando de lado el análisis de la variancia, consistió en plantear una estructura donde se variaba, simultáneamente, más de un factor, es decir, el diseño factorial y en calcular los distintos efectos factoriales. Así, frente

a un diseño de una sola dimensión de variación, propuso el diseño factorial de dos o más dimensiones de variación.

En el diseño clásico, cabe tener en cuenta dos aspectos que determinan su clasificación: la naturaleza de las variables independientes y el tipo de relación que se establece entre ellas. Cuando todas las variables independientes son manipuladas por el investigador, entonces se trata de un diseño factorial completamente al azar. Si, por otra parte, una de las variables independientes se utiliza como variable de clasificación, para formar submuestras de sujetos relativamente homogéneas con el propósito de reducir el error, se tiene un diseño de bloques aleatorizados. Como es obvio, la técnica de crear submuestras de sujetos más homogéneas, puede utilizarse con dos o más variables de clasificación, e incluso puede llegarse a soluciones más económicas, tales como los diseños de Cuadrado Latino, los diseños de Cuadrado Greco-latino, etc. Por último, con el propósito de optimizar el diseño y de reducir al máximo la variación debida al error, es posible sustituir los bloques por sujetos, con lo que se obtiene una estructura de medidas repetidas. El análisis de esta estructura, basada en el modelo mixto de análisis de la variancia, se halla por primera vez en el análisis del diseño *split-plot* (Hoshmand, 1994; Yandell, 1997) propuesto por Yates (1953).

El segundo aspecto, relativo a la estructura del diseño, se refiere a la clase de relación que se establece entre los factores. Esta relación puede ser de dos tipos, relación multiplicativa (cruzamiento) y relación de anidamiento (subordinación). La relación multiplicativa consiste en combinar todos los valores de un factor con todos los valores de los restantes factores; en cambio, la relación de anidamiento implica que los valores o los niveles de un factor no se cruzan con los de los restantes factores. Así, en los diseños factoriales, los grupos se derivan de las combinaciones entre los tratamientos experimentales o los valores de las distintas variables independientes. Por el contrario, en los diseños anidados, los valores de una primera variable independiente se aplican a dos o más grupos de sujetos y, a su vez, cada grupo recibe un valor distinto de una segunda variable de tratamiento o independiente. En este caso, la segunda variable es conocida como variable de tratamiento anidada. La diferencia más importante entre la relación de cruzamiento y de anidamiento es que, en la primera, cabe la posibilidad de estudiar el efecto combinado de las variables independientes o su interacción, mientras que, en la segunda, es imposible estimar el efecto cruzado.

Por último cabe destacar que el diseño clásico ha estado asociado, históricamente, al modelo de análisis de la variancia de modo que, como ha señalado Fisher (1935), estructura de análisis y modelo de análisis constituyen dos aspectos de una misma realidad. De ahí la necesidad de estudiar el diseño experimental clásico junto con el correspondiente análisis de datos. Este es el objetivo básico del presente manual, donde se describen las principales estructuras de diseño con su correspondiente procedimiento de análisis. No obstante, se ha asistido, a lo largo de los últimos años, a un drástico incremento de los diseños experimentales multivariados, debido a la disponibilidad de programas computacionales que han simplificado enormemente los cálculos estadísticos. El diseño experimental multivariado es una extensión de los diseños experimentales clásicos de una sola variable dependiente, propuestos por Fisher (1935). Así, los diseños multivariados incluyen dos o más variables dependientes que se registran, de forma simultánea, en los sujetos de la muestra. La principal ventaja de este enfoque es que, en un mismo diseño, se analizan relaciones complejas entre las medidas, así como el efecto que ejercen los tratamientos sobre el conjunto de tales medidas. Por el contrario, el modelo estadístico consiste en el análisis multivariante de la variancia y en el cálculo de las matrices de sumas cuadráticas y productos cruzados lo cual, como es obvio, supone un mayor grado de dificultad y complejidad (Bray y Maxwell, 1993).

Paralelamente a la tradición clásica, el diseño experimental ha seguido el enfoque centrado en el sujeto único. De ahí la notación que se ha utilizado para simbolizar el enfoque de grupos o de comparación de grupos, $\tilde{N}$ 1, y el enfoque de sujeto único, $\hat{N}$ 1 (Kratochwill y Levin, 1992). Este segundo enfoque, dado su interés aplicado, suele estudiarse como una estrategia de diseño particular, propia de contextos fundamentalmente educativos y clínicos. Cabe tener en cuenta que, si bien la experimentación de sujeto único entronca con los orígenes de la psicología científica, fue Skinner (1938) quien la adoptó de forma sistemática, y Sidman (1960) quien se encargó de dotar de una consistencia formal a esta nueva metodología, conocida originalmente como operante y que ha sido denominada posteriormente conductual o de sujeto único (Hersen y Barlow, 1976). No es nuestro propósito describir la estructura de estos nuevos diseños, donde las observaciones se toman de forma secuencial a lo largo de distintos períodos de tratamiento.

INVESTIGACIÓN EXPERIMENTAL Y DISEÑO

Un *experimento* es un plan estructurado de acción que, en función de una serie de objetivos, aporta la información necesaria para la prueba de expectativas teóricas o de hipótesis. Se trata de una guía que orienta los pasos del investigador, no sólo en la obtención de datos, sino en el correcto análisis y posterior inferencia de la hipótesis. De hecho, Fisher (1935) avanzó este concepto de experimento al afirmar que es una experiencia cuidadosamente planificada de antemano. Ello implica un conjunto de decisiones acerca de los tratamientos, del sistema de medida de la variable de respuesta, del control de los factores extraños, de la selección de los sujetos, del procedimiento, etc. (Brown y Melamed, 1993).

La experimentación se caracteriza, fundamentalmente, por la asignación totalmente aleatoria de los sujetos o de las unidades de observación a los distintos grupos o niveles de la variable de tratamiento. Los diseños experimentales en los que solo actúa el azar se denominan diseños completamente aleatorios. Ahora bien, el azar no siempre es la mejor garantía de precisión y eficacia en la inferencia de la hipótesis estadística. Por esa razón, tal y como se ha señalado anteriormente, a fin de extremar la precisión y la potencia, se utilizan distintas estructuras de diseño, tales como el diseño de bloques o el diseño de medidas repetidas. Las distintas versiones del diseño experimental se distinguen en función del concepto de eficacia o potencia de la prueba estadística, el cual requiere la reducción del tamaño del término de error, mediante estructuras que permitan controlar las fuentes de variación extrañas.

Una última cuestión sobre el diseño experimental hace referencia a su selección. Según Brown y Melamed (1993), en la selección del diseño experimental deben tenerse en cuenta las siguientes consideraciones: 1) la cantidad de variables independientes, 2) el origen y la cantidad de variables extrañas, 3) la cantidad de sujetos disponibles para la experiencia, y 4) cuestiones previas sobre las variables independientes. A modo de resumen, presentamos un diagrama de flujo para la selección del diseño en experimentos de uso frecuente, en función de la cantidad y de la modalidad de los grupos de los que se dispone (véase la Figura 5). Cabe señalar que en este diagrama se expone una clasificación elemental, que se amplía en el Capítulo 4 del presente manual, donde se proporciona una clasificación exhaustiva de los modelos de diseño experimental clásico que se utilizan con mayor frecuencia en el ámbito de las ciencias del comportamiento, en función de distintos criterios.

Nótese que los diseños experimentales suelen dividirse inicialmente en diseños de dos grupos y diseños de más de dos grupos. Los de dos grupos, según sean independientes o estén relacionados, determinan los diseños clásicos de dos grupos independientes o de dos

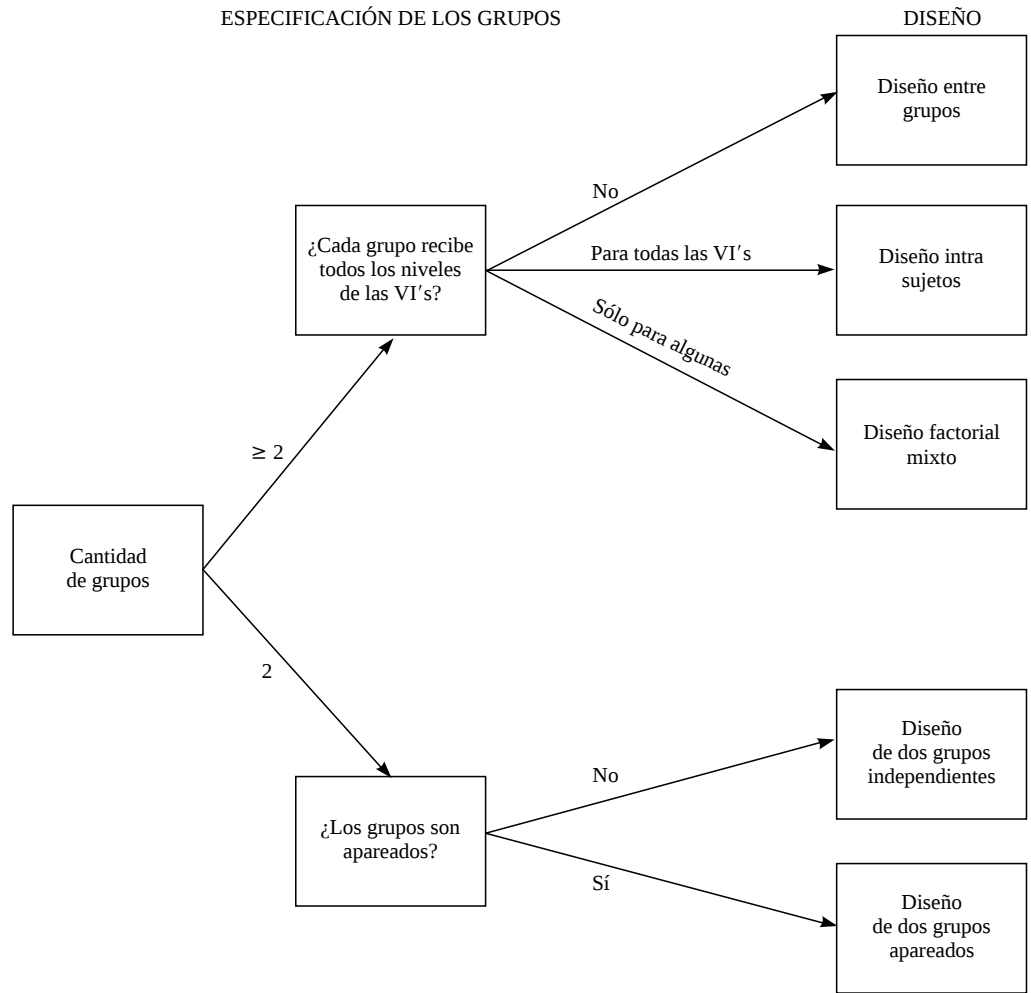


Figura 5 Selección del diseño experimental en función de la cantidad y modalidad de los grupos.

grupos apareados. Con experimentos de más de dos grupos cabe plantear si los grupos reciben tratamientos distintos, si un solo grupo de sujetos repite medidas, o bien se produce una combinación entre ambos enfoques, como en los diseños mixtos. Para cada uno de estos casos, es necesario definir el modelo de análisis o la prueba estadística en función de la estructura del diseño y de la naturaleza de la variable dependiente.

A MODO DE CONCLUSIÓN

A lo largo de la presente introducción se han descrito los aspectos esenciales de la investigación comportamental y se han proporcionado una serie de criterios para ubicar el diseño, tanto experimental como cuasi-experimental, en dicho ámbito de investigación. Ahora bien, dado que el texto se centra en el diseño experimental, se han expuesto las principales

líneas de desarrollo que ha seguido la investigación experimental a través del devenir histórico de la ciencia psicológica.

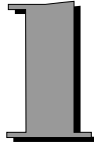
Considero que el lector interesado en el tema de la investigación y del diseño experimental va a encontrar, a lo largo de las páginas de este texto, un tratamiento completo y exhaustivo de las principales modalidades del diseño experimental clásico, complementado con ejemplos prácticos y su correspondiente tratamiento estadístico mediante el programa SPSS. Se trata, a mi entender, de un excelente manual de diseño experimental, que permitirá al estudiante hacerse con un conjunto de conceptos claros y precisos acerca del tema.

JAUME ARNAU Y GRAS
Catedrático de Psicología
Universidad de Barcelona

REFERENCIAS

- Arnau, J. (1990). Metodología experimental. En J. Arnau, M. T. Anguera y J. Gómez. *Metodología de la investigación en ciencias del comportamiento*. Murcia: Secretariado de Publicaciones de la Universidad de Murcia.
- Arnau, J. (1995). *Diseños experimentales en esquemas*. Barcelona: Publicacions Universitat de Barcelona.
- Arnau, J. (1997). *Diseños de investigación aplicados en esquemas*. Barcelona: Publicacions Universitat de Barcelona.
- Arnau, J. y Balluerka, N. (1998). *La psicología como ciencia: principales cambios paradigmáticos y metodológicos*. San Sebastián: Erein.
- Bray, J. H. y Maxwell, S. E. (1993). Multivariate analysis of variance. En Lewis-Beck, M. S. (Ed.), *Experimental design & methods*. Thousand Oaks, CA: Sage.
- Brown, S. R. y Melamed, L. E. (1993). Experimental design and analysis. En Lewis-Beck, M. S. (Ed.), *Experimental design & methods*. Thousand Oaks, CA: Sage.
- Campbell, D. T. y Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. En N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally. Reimpreso en 1966 como *Experimental and quasi-experimental design for research*. Chicago: Rand McNally.
- Campbell, D. T. y Boruch, R. F. (1975). Making the case for randomized to treatments by considering the alternatives: six ways in which quasi-experiments evaluations in compensatory education tend to underestimate effects. En C. A. Bennett y A. A. Lumsdaine (Eds.), *Evaluation and experiment*. Nueva York: Academic Press.
- Chow, S. C. y Liu, J. P. (1998). *Design and analysis of clinical trials. Concepts and methodology*. Nueva York: John Wiley.
- Clark-Carter, D. (1998). *Doing quantitative psychological research. From design to report*. Hove, East Sussex: Psychological Press.
- Cook, T. D. y Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.
- Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh*, 52, 399-433.
- Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgo: Oliver and Boyd.
- Fisher, R. A. (1935). *The design of experiments*. Londres: Oliver and Boyd.

- Gad, S. C. (1999). *Statistics and experimental design for toxicologists*. 3ra. ed. Boca Raton, FL: CRC Press.
- Hedrick, T. E., Bickman, L. y Rog, D. J. (1993). *Applied research design. A practical guide*. Newbury Park, CA: Sage.
- Hersen, M. y Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behaviour change*. Nueva York: Pergamon.
- Hoshmand, A. R. (1994). *Experimental research design and analysis*. Ann Arbor, CA: CRC Press.
- Kratochwill, T. R. y Levin, J. R. (1992). *Single-case research design and analysis*. Hillsdale. Nueva Jersey: Lawrence Erlbaum Associates.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Nueva York: Holt, Rinehart y Winston.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (ed. rev.). Chicago: University of Chicago Press.
- Madsen, K. B. (1978). Motivation and goals for global society. En E. Lazlo y J. Bierman (Eds.), *Studies on the conceptual foundations*. Nueva York: Pergamon Press.
- Madsen, K. B. (1980). Theories about history of sciences. *22 International Congress of Psychology*, Leipzig.
- Pedhazur, E. J. y Schmelkin, L. P. (1991). *Measurement, design, and analysis. An integrated approach*. Hillsdale, Nueva Jersey: Lawrence Erlbaum Associates.
- Petersen, R. G. (1985). *Design and analysis of experiments*. Nueva York: Marcel Dekker, Inc.
- Ross, A. S. y Grant, M. (1994). *Experimental and nonexperimental designs in social psychology*. Oxford: Westview Press.
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Nueva York: Basic Books.
- Skinner, B. F. (1938). *The behaviour of organisms*. Nueva York: Appleton-Century-Crofts.
- Yandell, B. S. (1977). *Practical data analysis for designed experiments*. Londres: Chapman & Hall.
- Yates, F. (1935). Discussion of Neyman's 1935 paper. *Journal of the Royal Statistical Society, Supplement 2*, 161-166.



CONCEPTO DE DISEÑO

El diseño es un concepto muy amplio, que se extiende a lo largo de varias etapas del proceso de investigación científica y que comprende tanto aspectos técnico-metodológicos como estadístico-analíticos.

Algunos autores definen el *diseño de investigación* como un conjunto de reglas a seguir para obtener observaciones sistemáticas y no contaminadas del fenómeno que constituye el objeto de nuestro estudio (García, 1992; Pereda, 1987). De acuerdo con esta acepción, Ander-Egg (1990), Arnau (1990b), Kerlinger (1979, 1986) y Martínez Arias (1983) categorizan el diseño como un plan estructurado de acción, elaborado en función de unos objetivos básicos y que se orienta a la obtención de datos relevantes para resolver el problema planteado. Moreno y López (1985) y Spector (1993) incluyen dentro de la actividad propia del diseño aspectos tales como la delimitación de problemas de investigación, el planteamiento de hipótesis operativas, la selección de variables, etc., dotando a este término de un significado muy amplio.

El concepto de «diseño» puede prestarse a múltiples caracterizaciones y definiciones en función del contexto en el que se utilice. En este sentido, López (1995) señala que lo que normalmente se entiende por diseño, en la tradición metodológica experimental, hace referencia a un conjunto de operaciones encaminadas a resolver un problema de investigación en términos causales, aunque, según el autor, la naturaleza y la cantidad de estas operaciones dista mucho de estar consensuada. A juicio de López, entre las operaciones características de un *diseño experimental* se deberían incluir:

- la operacionalización de hipótesis causales;
- la especificación de todas aquellas variables que pueden ejercer efectos sistemáticos sobre la conducta (es decir, de las variables independientes y perturbadoras);
- la especificación del número de unidades experimentales y de la población de la que éstas son extraídas;

- la delimitación del criterio de asignación de los tratamientos a las unidades experimentales;
- la especificación de las variables dependientes;
- la elección de las técnicas de análisis estadístico que se utilizan en la investigación.

En nuestra opinión, este conjunto de operaciones caracteriza, de forma muy adecuada, el concepto de diseño experimental. Dicha caracterización es compartida por una gran cantidad de autores entre los que cabe citar a Arnau (1986), Ato (1991), Keppel y Zedeck (1989), Maxwell y Delaney (1990), Mead (1988) y Milliken y Johnson (1984).

De acuerdo con la afirmación que realiza López (1995) acerca de la falta de consenso en la cantidad de operaciones que caracterizan un diseño experimental, encontramos diferentes conceptualizaciones que se distinguen fundamentalmente en dicho aspecto. Así, Robinson (1976) afirma que, en su sentido más general, el diseño experimental incluye las actividades básicas necesarias para efectuar correctamente un experimento, es decir, todas aquellas tareas que lleva a cabo el investigador desde la formulación de la hipótesis hasta la obtención de conclusiones. Dos de las cuatro acepciones que, según Meyers y Grossen (1974), puede tomar el término «diseño experimental» también le confieren un significado muy amplio. De hecho, utilizado como verbo, el vocablo *diseñar* hace referencia a un proceso unitario de actividades y de decisiones encaminadas a planificar una investigación. Por otra parte, cuando el término *diseño* se utiliza como sustantivo, hace alusión a un determinado tipo de procedimiento que se emplea para organizar los diferentes aspectos y materiales necesarios para probar un efecto determinado. En otras palabras, describe una clase específica de experimento.

Una segunda línea en la caracterización del diseño experimental restringe sus actividades a aquellas operaciones de carácter técnico encaminadas a un fin. Según Arnau (1986), esta conceptualización es la más común entre los teóricos del diseño. Por lo general, las actividades que configuran el diseño, desde esta perspectiva, hacen referencia a las dos restantes acepciones planteadas por Meyers y Grossen (1974), a saber, al modo en que se forman los diferentes grupos experimentales y a la técnica de análisis estadístico que se aplica a los datos. De acuerdo con esta conceptualización, Kirk (1982) considera el término *diseño experimental* como sinónimo de un plan para la asignación de las condiciones experimentales a los sujetos y del análisis estadístico asociado con dicho plan. De forma similar, Arnau (1994) define el *diseño experimental* como un plan de investigación mediante el que se pretende contrastar el efecto causal de una o más variables manipuladas por el experimentador, e incluye dentro de este plan el procedimiento de asignación de los sujetos a los distintos niveles de tratamiento y la selección de una adecuada técnica de análisis. Esta caracterización, a la que nosotros nos sumamos, es también compartida por autores como Edwards (1985), Keppel (1982) y Winer, Brown y Michels (1991), entre otros.

En relación con el *objetivo del diseño experimental*, la mayoría de los autores coinciden en afirmar que su propósito principal consiste en detectar, de manera inequívoca, qué influencia ejerce(n) la(s) variable(s) independiente(s) sobre la(s) variable(s) dependiente(s). Arnau (1994), por ejemplo, opina que el objetivo esencial del diseño radica en comprobar los posibles efectos de una o más variables de tratamiento sobre uno o más indicadores conductuales. El autor añade que en el razonamiento lógico subyacente a la prueba de la hipótesis de causalidad, se asume que las diferencias observadas entre las medias de los grupos en la(s) variable(s) dependiente(s) se deben única y exclusivamente a los cambios producidos en la(s) variable(s) manipulada(s). La ya clásica definición de O'Neil (1962) del término «diseño experimental» guarda una estrecha relación con tal objetivo. De hecho, el autor

conceptualiza el diseño como un modelo particular de variación y de constancia. De variación, porque el experimentador manipula o hace variar sistemáticamente la variable independiente, y de constancia, porque mantiene constantes el resto de variables. En tales circunstancias, los cambios observados en la(s) variable(s) dependiente(s) pueden atribuirse de forma inequívoca a la manipulación de la(s) variable(s) independiente(s).

2

PRINCIPALES ALTERNATIVAS METODOLÓGICAS Y DISEÑOS EN PSICOLOGÍA: UNA PERSPECTIVA GENERAL

2.1. PRINCIPALES ESTRATEGIAS PARA LA RECOGIDA DE LOS DATOS

El proceso de investigación científica participa de tres niveles de actuación que son comunes a cualquier objeto de estudio (Arnau, 1989, 1990b). Tales niveles son el *nivel teórico-conceptual*, el *nivel técnico-metodológico* y el *nivel estadístico-analítico*. El nivel teórico-conceptual hace referencia a la primera etapa del método científico, en la que se elaboran las representaciones abstractas de los fenómenos reales mediante el planteamiento del problema, la formulación de las hipótesis y la discusión de los resultados. Por su parte, en el nivel estadístico-analítico se lleva a cabo el análisis estadístico de los datos obtenidos a partir del diseño, a fin de verificar o rechazar las hipótesis propuestas.

A nuestro juicio, el segundo nivel constituye la etapa más importante dentro de dicho proceso. El nivel técnico-metodológico se caracteriza por dos actividades básicas: la *operacionalización de la hipótesis* y la *elección del diseño de investigación*. Así, requiere, entre otras cosas, que el investigador seleccione los indicadores más adecuados para representar los constructos teóricos incluidos en la hipótesis y decida cuál es la estrategia más apropiada para obtener los datos que le permitan verificarla. Es decir, debe decidir qué *tipo de metodología* y, por ende, qué *tipo de diseño* va a emplear en el transcurso de la investigación.

En esta etapa, el método general de investigación se diversifica en tres formas de actuación científica: la *metodología experimental*, la *metodología selectiva* y la *metodología observacional* (Arnau, Anguera y Gómez, 1990).

Esta clasificación se fundamenta en el grado de manipulación o de control interno ejercido sobre la situación de observación (Anguera, 1990, 1991a, 1991b; Arnau, 1978a, 1990b; Ato, 1991; Gómez, 1990; Martínez Arias, 1983, 1986). Así, los métodos quedarían localizados en un eje continuo que se desplaza desde un máximo control o intervención por parte del investigador (*metodología experimental*) hasta una presencia mínima de manipulación y control en la situación de observación (*metodología observacional* o *naturalista*), pasando

por un nivel intermedio, en el que control y naturalidad se combinan de forma complementaria (*metodología selectiva o de encuesta*).

De acuerdo con esta categorización, Ato (1991) señala que en la investigación de fenómenos comportamentales pueden distinguirse tres compromisos básicos: el *realismo* de las variables, la *aleatorización* de los tratamientos y la *representatividad* de las unidades de muestreo respecto de la población básica de referencia. Según el autor, si el énfasis se centra en el criterio del *realismo*, la investigación se fundamenta sobre la *metodología observacional*. Si se centra en el criterio de la *representatividad*, prevalece la *metodología de encuesta*, y si se centra en el criterio de la *aleatorización*, la investigación descansa sobre la *metodología experimental*.

La distinción propuesta por Kish (1987) entre estas tres metodologías también gira en torno a tres criterios, dos de los cuales coinciden plenamente con los planteados por Ato. Tales criterios son el *control*, la *representatividad* y el *realismo*.

La **metodología experimental** se caracteriza por el *control* establecido en la situación de observación. Siguiendo al autor, el término *control* hace referencia al proceso de asignación aleatoria de las unidades experimentales a las diferentes condiciones de tratamiento, a la manipulación o variación controlada de la variable independiente y a las técnicas específicas utilizadas para evitar la influencia de variables contaminadoras o extrañas. En definitiva, se trata de una estrategia en la que los datos se obtienen creando las condiciones específicas para que se produzcan los fenómenos que son objeto de estudio (Arnau, 1994).

De esta forma, la metodología experimental se identifica con la búsqueda de *relaciones causales*, ya que permite satisfacer los tres supuestos exigibles a cualquier relación para poder ser considerada causal (Kenny, 1979; Suppes, 1970), a saber, la precedencia temporal de la causa, la covariación entre causa y efecto y la ausencia de espuriedad.

La **metodología de encuesta** se caracteriza por la *representatividad* de la información obtenida. Así, la preocupación principal del investigador que emplea esta metodología se centra en la selección de una *muestra representativa*, ya que el objetivo básico de los estudios selectivos consiste en generalizar los datos obtenidos a partir de una muestra, a la población de la que se extrae dicha muestra. Las *encuestas* y los *estudios correlacionales* son aplicaciones concretas de este criterio de representatividad.

El *realismo* es la condición esencial que caracteriza a la **metodología observacional**. Cuando utiliza esta estrategia, el investigador registra la información en un marco natural, observando sistemáticamente a un sujeto o a un grupo de sujetos, sin realizar ningún tipo de intervención sobre la situación observada.

Como señala López (1995), cada una de estas tres metodologías maximiza una característica de los datos obtenidos que, aunque en menor medida, no quedan exentos de la influencia de las otras dos estrategias.

En la Tabla 2.1 se resumen los principales aspectos, expuestos hasta el momento, que definen la actuación científica.

TABLA 2.1 Criterios de clasificación de la actuación científica

Metodología	Control	Compromiso
Experimental	Máximo	Aleatorización
Selectiva o de encuesta	Combinación de control y naturalidad	Representatividad
Observacional o naturalista	Mínimo	Realismo

Además de presenciar la relevancia adquirida por la clasificación que acabamos de abordar, la historia de la psicología ha sido y sigue siendo testigo de múltiples categorizaciones en torno a los diferentes métodos o estrategias de investigación que pueden emplearse en dicha disciplina. A continuación recogemos algunas de tales clasificaciones.

Las primeras discusiones sobre los métodos de la psicología fueron provocadas por Cronbach (1957), quien planteó la dicotomía entre *psicología experimental* y *psicología correlacional*. A este respecto es importante subrayar que, si bien ha constituido una constante fuente de confusión, por *estudio correlacional* no entendemos aquel en cuyo nivel analítico se emplea la correlación como técnica estadística, sino que hacemos referencia a una forma global de proceder al llevar a cabo la investigación. Incidiendo en sus principales características, Cronbach (1957) afirma que la *psicología experimental* se centra en el análisis de los procesos psicológicos básicos, utilizando para ello métodos de investigación de extremo control. La *psicología correlacional*, por su parte, se caracteriza por la aplicación de métodos de análisis multivariado para el estudio de las diferencias individuales. Campbell y Stanley (1963, 1988; véase también Cook y Campbell, 1979; Cook, Campbell y Peracchio, 1990) rompen esta dicotomía al introducir entre ambos polos una tercera categoría de modelos de investigación: los denominados *diseños cuasi-experimentales*. Como veremos posteriormente, la clasificación tripartita defendida por estos autores se basa en los criterios de *manipulación* y *aleatorización*. Así, las investigaciones que cumplen con ambos criterios se consideran *experimentales*. Aquellas que sólo responden al primer requisito se categorizan como *cuasi-experimentales*. Y, por último, las investigaciones se definen como *correlacionales* cuando no cumplen con ninguno de estos dos principios.

Otra de las clasificaciones frecuentemente utilizadas en el ámbito de las ciencias del comportamiento es la que distingue entre *métodos cuantitativos* y *métodos cualitativos*. En la concepción cuantitativa de la ciencia, el objetivo de la investigación consiste en establecer relaciones causales que permitan explicar determinados fenómenos, siguiendo para ello una perspectiva *nomotética*. Por el contrario, la interpretación de los fenómenos que se examinan utilizando métodos cualitativos es de naturaleza *ideográfica*. Aunque ambas metodologías parten de concepciones diferentes con respecto a la ciencia, Cook y Reichardt (1982) destacan que, desde el punto de vista metodológico, sus aportaciones no deben considerarse mutuamente excluyentes, sino complementarias.

Festinger y Katz (1953) plantean una categorización que también ha llegado a tener muchos adeptos en el ámbito de la psicología. En ella, los autores establecen una distinción entre *experimentos de laboratorio*, *experimentos de campo* y *estudios de campo*. El criterio de delimitación entre tales métodos consiste en el diferente grado de control que puede ejercer el investigador sobre la situación de observación. Así, mientras en los *experimentos de laboratorio* el control adquiere su máxima expresión, en los *estudios de campo* el investigador se limita únicamente a observar. Partiendo de un menor grado de control que el existente en el laboratorio, los *experimentos de campo* son experimentos que se realizan en situaciones de la vida real. Los *estudios de campo* se asemejan a los *métodos selectivos* o *de encuesta* cuando utilizan cuestionarios o entrevistas para la recogida de los datos. Sin embargo, cabe señalar que mientras en los *métodos de encuesta* se emplean muestras representativas para la recogida de datos y se infieren los procesos psicológicos y sociales a partir del producto final o del resultado obtenido por el sujeto en la encuesta, en los *estudios de campo* tales procesos se observan y se registran durante su desarrollo.

En estrecha relación con la metodología utilizada, el investigador debe tomar una serie de decisiones respecto a la forma concreta en la que han de obtenerse los datos, es decir, tiene que plantear el **diseño de la investigación**. Cuando el diseño es conceptualizado

en función de los objetivos que se persiguen en el estudio, Arnau (1995a) establece una distinción entre *diseños experimentales*, *diseños cuasi-experimentales* y *diseños no-experimentales*, incluyendo en esta última categoría los *diseños de encuesta* y los *diseños observacionales*. Por otra parte, cuando se adopta como criterio de clasificación el procedimiento utilizado para la obtención de los datos, el autor distingue entre los *diseños transversales* y los *diseños longitudinales*. Los *primeros* se aplican en aquellas situaciones en las que se tiene en cuenta un solo registro por período de observación y unidad. Los *diseños longitudinales*, por el contrario, se caracterizan por registrar un conjunto de medidas de las mismas unidades observacionales en distintos períodos temporales. Obviamente, ambos criterios no son excluyentes.

Según Arnau (1994), los *diseños experimentales* y, en menor medida, los *diseños cuasi-experimentales* se ajustan a la *estrategia* o al *paradigma experimental*. Los *diseños de encuesta* y los *diseños observacionales*, por su parte, se integran bajo la *estrategia no-experimental*. En las siguientes líneas abordamos estos cuatro tipos de diseños.

2.2. DISEÑOS EXPERIMENTALES

Spector (1993) describe el *diseño experimental* como una combinación entre los conceptos «constancia», «comparación», «aleatorización» y «control». En este tipo de estructura de investigación algunas variables se comparan entre sí, otras se mantienen constantes a un determinado nivel y, por tanto, se controlan, y otras pueden variar sin restricción alguna bajo el supuesto de que sus posibles efectos perturbadores son promediados gracias al azar. El diseño experimental típico se aplica en el laboratorio y posee dos características distintivas esenciales: (1) el *control* o *manipulación activa* de una o más variables independientes y (2) la utilización de una *regla de asignación aleatoria* para asignar los sujetos a las condiciones de tratamiento y, en el caso de disponer de un solo sujeto, para asignar las diferentes condiciones o tratamientos a dicho sujeto.

La primera característica implica que el experimentador selecciona deliberadamente los valores de la variable independiente y, además, crea las condiciones necesarias para la producción artificial de tales valores. En otras palabras, el experimentador interviene activamente sobre las condiciones que se supone que son la causa de los cambios constatados en la conducta de los sujetos.

En relación con la segunda característica, cabe decir que se considera el elemento esencial del diseño experimental clásico, ya que posibilita que la variable de asignación no correlacione, dentro de límites probabilísticos, con ninguna otra variable. De esta forma, se garantiza la equivalencia entre los diferentes grupos experimentales en un gran conjunto de variables antes de administrar los tratamientos.

Otra de las características especialmente importantes del diseño experimental es su capacidad para controlar adecuadamente las posibles variables perturbadoras y, por tanto, su alto grado de *validez interna*. En virtud de la aleatorización, los diseños experimentales permiten un control preciso de toda fuente de variación alternativa a la que pueden atribuirse las diferencias entre los grupos.

Respecto a la *inferencia de la hipótesis*, el diseño experimental es la única estructura de trabajo que garantiza plenamente la inferencia de *hipótesis explicativas* (Arnau, 1995a). La *inferencia conceptual* se caracteriza por el proceso de extraer contenidos significativos de acuerdo con los propósitos de la investigación. Como afirman Pedhazur y Pedhazur Schmelkin (1991), en la investigación experimental y cuasi-experimental, las inferencias se realizan

desde la(s) variable(s) independiente(s), o supuesta(s) causa(s), a la(s) variable(s) dependiente(s) o fenómeno(s) objeto de explicación. En este sentido, cuando se manipula la variable independiente, se asume que dicha variable es la causa de los cambios operados en la variable dependiente y esta hipótesis o predicción es objeto de verificación e inferencia dentro de las estrategias experimental y cuasi-experimental. Por tanto, el diseño experimental permite establecer *relaciones de concomitancia o causalidad*.

En concordancia con esta lógica inferencial, Arnau (1995a) señala que en los diseños experimentales las *comparaciones* se realizan *entre los grupos*, ya que, debido a la aleatorización, éstos son equivalentes en todos aquellos aspectos diferentes a la variable de tratamiento. En consecuencia, cualquier diferencia existente entre tales grupos puede atribuirse, de forma inequívoca, a la acción de la(s) variable(s) manipulada(s).

2.3. DISEÑOS CUASI-EXPERIMENTALES

Los *diseños cuasi-experimentales* juegan un papel primordial en los contextos de investigación aplicada. Normalmente, el objetivo de estos diseños consiste en comprobar el efecto de determinados tratamientos terapéuticos o programas de intervención social o educativa. En este sentido, el *cuasi-experimento*, como modelo de investigación derivado del paradigma experimental, se caracteriza por el estudio de la variable de tratamiento en contextos donde el investigador no puede asignar las unidades de análisis a los distintos niveles de la(s) variable(s) de interés.

Cook (1983) define los cuasi-experimentos como una clase de estudios empíricos a los que les faltan algunos de los rasgos usuales de la experimentación. Habitualmente, se llevan a cabo fuera del laboratorio y no implican asignación aleatoria de las unidades experimentales a las condiciones de tratamiento. Sin embargo, al igual que en el caso de los diseños experimentales, estos diseños pretenden establecer relaciones de causalidad entre la(s) variable(s) independiente(s) y la(s) variable(s) dependiente(s). Además, su estructura implica tanto la manipulación de una o más variables independientes como la medida de la(s) variable(s) dependiente(s).

En relación con la *inferencia de hipótesis*, es preciso señalar que, si bien la inferencia causal es posible dentro del ámbito cuasi-experimental, la cantidad de *hipótesis explicativas rivales* que pueden competir con la hipótesis planteada por el investigador es muy alta. Ello se debe a que en los diseños cuasi-experimentales se carece de un adecuado control experimental sobre las posibles fuentes de confundido, de ahí que el investigador se enfrente con la difícil tarea de tener que separar los efectos de la variable independiente de la influencia de cualquier otra variable extraña. En consecuencia, como señalan Dowdy y Wearden (1991), los cuasi-experimentos poseen menor potencia inferencial que los experimentos.

En lo que respecta al tipo de comparaciones que se realizan en tales diseños, Arnau (1995a) afirma que se mantiene el criterio de *comparación de grupos* como elemento primordial para la inferencia de la hipótesis, aunque, como ya se ha señalado, la falta de aleatorización impide asegurar la exclusión de factores extraños de confundido.

Pedhazur y Pedhazur Schmelkin (1991) proporcionan una figura muy ilustrativa para distinguir los *experimentos* de los *cuasi-experimentos* (véase la Figura 2.1).

En la sección (a) se representa un experimento, donde la variable X o variable independiente ejerce influencia sobre la variable Y o variable dependiente. Dado que los sujetos han sido asignados al azar a los distintos valores de X cabe esperar que, en términos probabilísticos, sean idénticos respecto de otras variables que pueden afectar a Y, variables que se

subsumen bajo el término de error (ε). En consecuencia, podemos suponer que no existe correlación entre X y ε . En la sección (b), la variable X también es la causa de la variable Y , sin embargo los valores de X pueden estar correlacionados con el término de error. Este fenómeno se conoce como el *problema de la tercera variable* y entorpece en gran medida el establecimiento de relaciones causales entre X e Y . La cuestión tiene solución si se conoce la regla de asignación de los sujetos, ya que en tal caso es posible ajustar estadísticamente los resultados sustrayendo de la varianza de Y la proporción de variación atribuible a la regla de asignación (Pascual, García y Frías, 1995).

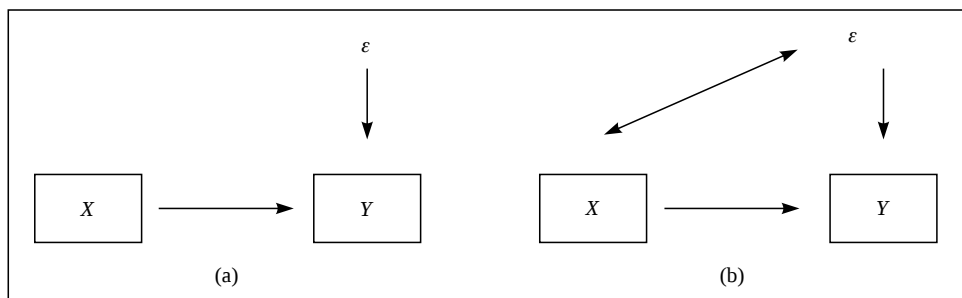


Figura 2.1 Relaciones entre la variable independiente, la variable dependiente y el término de error en el experimento y en el cuasi-experimento.

2.4. DISEÑOS NO-EXPERIMENTALES

En esta categoría se incluyen los *diseños de encuesta* y los *diseños observacionales*.

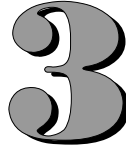
El objetivo principal de los **diseños de encuesta** consiste en la descripción de las características o propiedades de una población, aunque también pueden tener otros propósitos tales como el estudio de los procesos de cambio y de las relaciones entre distintas variables. Estos diseños se caracterizan por estar basados en muestras de individuos seleccionadas al azar de entre una o más poblaciones, no siendo factible la asignación aleatoria de los sujetos a los distintos niveles de la(s) variable(s) independiente(s) (Gómez, 1990). Al igual que los estudios observacionales, se limitan a la medición empírica de características y conductas de individuos o grupos tal y como éstos aparecen en su contexto, diferenciándose de la observación, en la posibilidad de inferir los valores muestrales a los poblacionales, debido a la selección aleatoria de los sujetos experimentales. En definitiva, como apunta Gómez (1990), los diseños de encuesta pretenden describir y analizar ciertos fenómenos sin manipular las variables que estudian, pero poniendo en evidencia distintos niveles de éstas mediante la adecuada selección de las muestras. Atendiendo a la forma en la que se administran las encuestas, tales diseños pueden utilizar como instrumento de recogida de datos la *entrevista* o el *cuestionario*.

Los **diseños observacionales** tienen como objetivo principal la descripción de los fenómenos que ocurren en ambientes naturales. También permiten estudiar los procesos de cambio y plantear cuestiones relativas a la relación entre variables, pero en condiciones de escaso control. Como señala Spector (1993), en tales diseños el control no implica la manipulación activa de las condiciones experimentales, sino la selección de aquellos casos o de aquellas variables que cumplen determinados criterios y que resultan de interés para el investigador. Estos diseños se caracterizan por el grado de interacción entre las unidades observacionales

y el investigador, así como por el grado de estructuración o control de la situación observada. Dichas estructuras de investigación utilizan dos amplios procedimientos de recogida de datos, a saber, las *técnicas de observación sistemática*, en las que el investigador no participa en la situación observada, y las *técnicas de observación participante*, las cuales requieren que el investigador forme parte de la situación que se observa.

Con respecto a la clase de hipótesis que permiten formular los diseños no-experimentales, Arnau (1995a) afirma que aunque en algunos casos posibilitan el planteamiento de teorías explicativas, normalmente responden a objetivos descriptivos y predictivos. Pedhazur y Pedhazur Schmelkin (1991) plantean que, a diferencia de lo que ocurre en la investigación experimental y cuasi-experimental, en estas estructuras de trabajo la *inferencia conceptual* parte de la(s) variable(s) dependiente(s) para intentar identificar o descubrir la(s) variable(s) independiente(s). Así, el investigador se abstiene de formular hipótesis a fin de no perder la objetividad y prefiere extraer inferencias a partir de la mera observación.

En relación con los *grupos experimentales*, cabe señalar que éstos se forman en función de la(s) variable(s) dependiente(s). En este caso, el investigador trata de averiguar qué es lo que ha ocasionado las diferencias observadas entre los grupos. Sin embargo, hay que destacar que, cuando los grupos se forman en función de los valores que presentan los sujetos en la(s) variable(s) dependiente(s) y las diferencias entre ellos se atribuyen a determinadas causas, resulta imposible excluir las *hipótesis alternativas* o *de efectos rivales*. En otras palabras, los diseños no-experimentales carecen de procedimientos adecuados para el control de las posibles fuentes de confundido.



ASPECTOS ESENCIALES DE LA METODOLOGÍA EXPERIMENTAL

3.1. DEFINICIÓN Y OBJETIVOS DE LA EXPERIMENTACIÓN

3.1.1. Paradigma experimental contra paradigma asociativo

Arnau (1995a) conceptualiza el vocablo *paradigma* como un sistema inspirador de las metodologías de trabajo y considera que, en psicología, se pueden distinguir dos paradigmas o tradiciones: el *paradigma experimental* y el *paradigma asociativo*. Cada uno de ellos se caracteriza por la formulación de un determinado tipo de hipótesis, por el grado en que interviene el investigador en la situación estudiada, por las técnicas de recogida de datos y por los procedimientos empleados para la verificación de las hipótesis.

Siguiendo a este autor, y retomando algunas de las ideas expuestas en el Capítulo 2, cabe señalar que mientras el paradigma experimental pretende establecer *hipótesis de carácter causal* y, por tanto, inferir teorías explicativas de naturaleza causal, el paradigma asociativo enfatiza las *hipótesis de covariación* y la confirmación de enunciados relacionales. De acuerdo con tales hipótesis, los diseños derivados del paradigma experimental son los *diseños experimentales* y *cuasi-experimentales*, mientras que los *diseños de encuesta* y los *diseños observacionales* se integran en el paradigma asociativo.

En segundo lugar, la tradición experimental se caracteriza por la *manipulación* de las condiciones, de los hechos o de las situaciones, es decir, por la intervención activa del investigador en el fenómeno que se somete a estudio. El paradigma asociativo, por el contrario, sólo requiere el registro simultáneo de dos o más variables, sin ejercer ningún tipo de manipulación sobre tales variables.

Además de manipular las condiciones de tratamiento, en la tradición experimental también se lleva a cabo un adecuado *control* de las variables extrañas o de los factores de confundido, de forma que se excluye la posibilidad de que una tercera variable pueda explicar los resultados obtenidos. Este control resulta inadecuado en la tradición asociativa.

Por último, en el paradigma experimental la verificación de las hipótesis se basa en una *evidencia de concomitancia*, de ahí que el método experimental permita establecer relaciones de carácter causal entre las variables. Lo que se infiere a partir de la tradición asociativa, por el contrario, son relaciones de cambio entre dichas variables, ya que en ella la verificación de las hipótesis se fundamenta en una *constatación de simultaneidad o contigüidad*.

TABLA 3.1 Paradigma experimental contra paradigma asociativo

	Paradigma experimental	Paradigma asociativo
Tipo de hipótesis	Causal	De covariación
Grado de intervención del investigador	Activa → Manipulación de condiciones	No manipulación: Registro simultáneo de variables
Control	Riguroso	No control
Verificación de hipótesis	Evidencia de concomitancia	Constatación de simultaneidad o contigüidad
Tipo de diseño	Experimental y Cuasi-experimental	Encuesta y Observacional

No obstante, cabe señalar que entre las dos estrategias de estudio propias del paradigma experimental, es decir, entre el *experimento* y el *cuasi-experimento*, la que aquí nos ocupa tiene mayor potencia inferencial. Ello es así porque además de poseer todas las características que definen el paradigma experimental, el experimento también se caracteriza por la presencia de *aleatorización* en el proceso de asignación de los sujetos a las condiciones de tratamiento y, como ya se ha apuntado en el epígrafe anterior, la aleatorización constituye un elemento decisivo en el control de las fuentes extrañas de variación.

3.1.2. Concepto de experimento

Zimny (1961) define el *experimento* como «una observación objetiva de fenómenos, a los cuales se les hace ocurrir bajo situaciones de estricto control y en los que se hacen variar uno o más factores, mientras los restantes permanecen constantes». Los elementos que caracterizan al experimento en esta definición, a saber, el *control* y la *manipulación*, también son considerados como sus rasgos esenciales por otros muchos autores inmersos en el ámbito de las ciencias del comportamiento (p. e. Cozby, 1993; Kantowitz, Roediger y Elmes, 1991; Kiess y Bloomquist, 1985; Reaves, 1992). De acuerdo con esta conceptualización, Fisher (1935), uno de los autores clásicos de mayor relevancia en este campo de trabajo, define el experimento como «una experiencia cuidadosamente planificada de antemano», dotando al mismo de un carácter claramente activo. A este respecto, Brown y Melamed (1993) señalan que el método activo consiste en manipular niveles de variables independientes (causas), seleccionadas con el objetivo de examinar la influencia que ejercen sobre determinadas variables dependientes (efectos).

No obstante, además de la manipulación y del control, existe otra característica que la mayoría de los autores consideran como requisito esencial para que un estudio pueda conceptualizarse como experimental: la *aleatorización*. De acuerdo con una gran cantidad

de autores contemporáneos (Christensen, 1988; Dunham, 1988; Levine y Parkinson, 1994; Shaughnessy y Zechmeister, 1994; Suter, Lindgren y Hiebert, 1989), Pedhazur y Pedhazur Schmelkin (1991) definen el experimento como «una investigación en la que al menos una de las variables es manipulada, y las unidades experimentales son asignadas aleatoriamente a los distintos niveles de la(s) variable(s) manipulada(s)». Partiendo de esta definición, Pascual, García y Frías (1995) consideran que lo que actualmente se entiende por *investigación experimental* en psicología y en otras ciencias sociales, requiere el cumplimiento de las siguientes condiciones, que se ilustran en la Figura 3.1:

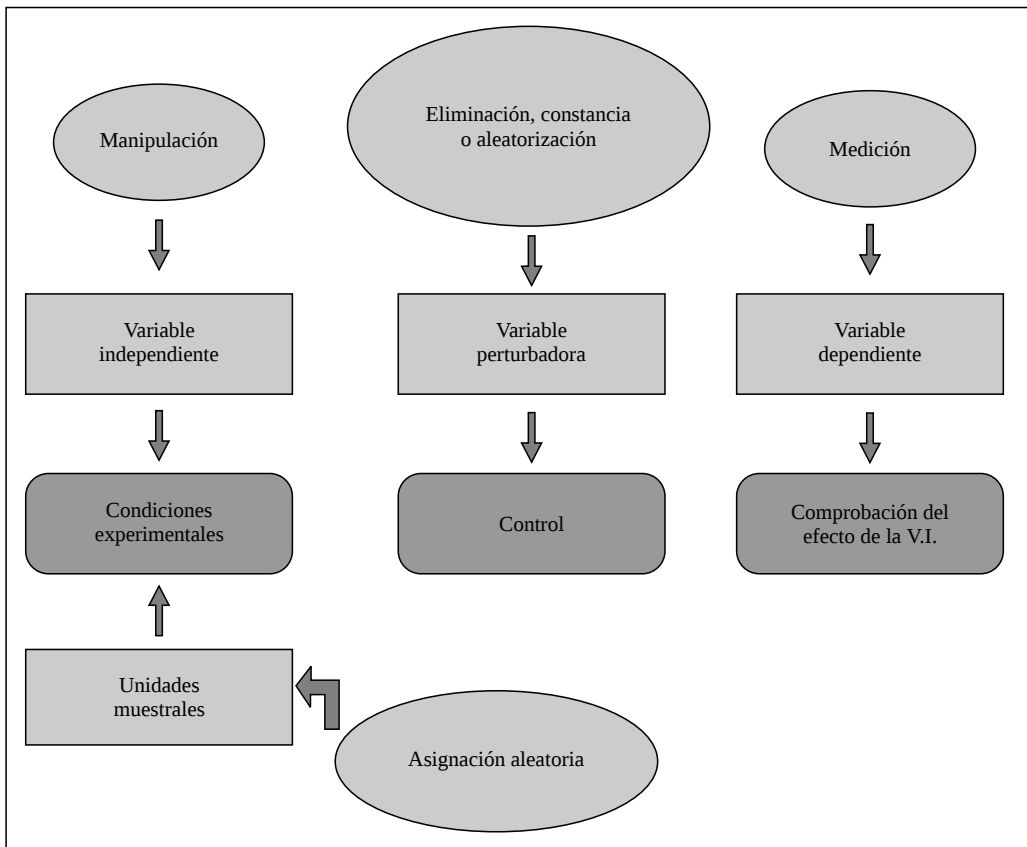


Figura 3.1 Condiciones que debe cumplir una investigación experimental.

1. La existencia de, al menos, una variable manipulada que se denomina *variable independiente*. Esta manipulación implica la selección de una serie de valores de dicha variable y la delimitación de una condición experimental distinta para cada uno de tales valores.
2. La asignación aleatoria de las unidades experimentales a las diferentes condiciones de tratamiento.
3. La comprobación del efecto que ejerce(n) la(s) variable(s) independiente(s) sobre determinada(s) conducta(s) o variable(s) de medida, que se conoce(n) como *variable(s) dependiente(s)*.

4. El control de cualquier otra variable que pueda ser fuente de variación de los datos y que no haya sido manipulada por el investigador. Este tipo de variables se denominan *variables extrañas* o *perturbadoras* y, si se quieren obtener resultados fiables, dichas variables deben ser eliminadas, mantenidas constantes o, simplemente, aleatorizadas.

Las variables extrañas no son siempre conocidas y, por ello, con frecuencia resultan difíciles de controlar mediante eliminación o constancia. Esta es una de las razones por las que el proceso que mejor garantiza el control experimental es la aleatorización. La asignación aleatoria de las unidades experimentales a los distintos niveles de la(s) variable(s) manipulada(s) permite asumir que los grupos son equivalentes con respecto a cualquier variable desconocida y, por tanto, comparables entre sí. Debido a la homogeneidad existente entre los grupos antes de recibir los tratamientos experimentales, cualquier diferencia que se constate tras el registro de las puntuaciones obtenidas en la(s) variable(s) dependiente(s) puede atribuirse de forma inequívoca a la manipulación de la(s) variable(s) independiente(s). Esta es la razón por la que algunos autores (p. e. Cook y Campbell, 1979; Dwyer, 1983) se refieren al diseño experimental como «experimento verdadero» en oposición al diseño cuasi-experimental, en el que la asignación aleatoria no se halla presente.

3.1.3. Objetivos del experimento

Como afirman Pascual, García y Frías (1995), el experimento persigue dos objetivos fundamentales: la comprobación de teorías y la estimación de los efectos producidos por uno o más factores sobre determinado(s) fenómeno(s).

En el ámbito de la investigación básica, el experimento de laboratorio se utiliza principalmente para **verificar enunciados teóricos**, es decir, para comprobar hipótesis o teorías. La función de toda ciencia, y entre ellas la de la psicología, consiste en comprender y en explicar los fenómenos que estudia. Aunque muchos de los fenómenos psicológicos son procesos, mecanismos o estructuras que no pueden observarse directamente, es posible proponer teorías acerca de su funcionamiento. Tales teorías adquieren el estatus de modelos científicos cuando se someten a comprobación empírica mediante la formulación de las hipótesis pertinentes. En definitiva, la psicología pretende descubrir los principios generales de la conducta, planteando diversas teorías sobre su funcionamiento y encontrando evidencia empírica que confirme dichas teorías.

Cuando el psicólogo experimental somete a comprobación una hipótesis o una implicación lógica derivada de una teoría, lo que fundamentalmente pretende no es establecer una relación entre la(s) variable(s) independiente(s) y la(s) variable(s) dependiente(s), sino obtener una evidencia en favor del supuesto teórico que ha formulado previamente, es decir, demostrar la veracidad del modelo teórico que ha planteado para explicar determinados aspectos de la realidad.

Las hipótesis que se formulan en el ámbito científico adoptan la forma de un enunciado condicional del tipo «si *a* entonces *b*». Para contrastar este enunciado se sigue el proceso que en lógica se conoce como *modus tollens*:

«Si se dan las condiciones *p*, entonces sucede *q*,
no sucede *q*,
luego *p* no es verdadero (y debe buscarse otra explicación)».

No obstante, dado que las hipótesis nunca aparecen aisladas, el *modus tollens* suele presentarse de la siguiente forma:

«Si p y q y r y y t , entonces c .»

Pero si se demuestra empíricamente que c no es verdadera, entonces p , q , r , t no son verdaderas (basta con que no lo sea una de ellas).

Por tanto, poner a prueba la hipótesis consiste en comprobar que cada vez que ocurre p , tiene lugar q . Si esta secuencia no se produce, se rechaza la hipótesis. Si, por el contrario, q ocurre, la hipótesis queda confirmada. Aunque no es posible afirmar con toda seguridad que la hipótesis es verdadera (ya que puede haber otras hipótesis que también son compatibles con los resultados), se puede concluir que existe evidencia a favor de ella.

El proceso que se sigue en la práctica para contrastar una hipótesis en el ámbito de la psicología requiere pasar por una serie de fases y llegar, finalmente, a la *prueba de la hipótesis estadística*. Recordemos que cabe distinguir entre dos tipos de hipótesis estadísticas: la *hipótesis de nulidad* y la *hipótesis alternativa*. En el modelo fisheriano, el investigador siempre parte del supuesto de que la hipótesis nula es verdadera, siendo este supuesto el que se somete a prueba. Si las diferencias encontradas en la(s) variable(s) dependiente(s) entre los sujetos sometidos a diferentes tratamientos tienen una baja probabilidad de aparición bajo el supuesto de que la hipótesis nula sea verdadera, se rechaza dicha hipótesis y se acepta la hipótesis alternativa como la explicación más plausible, aunque siempre debe asumirse un determinado riesgo de error en la decisión.

Por otra parte, en el ámbito de la investigación aplicada, el principal objetivo del experimento no consiste en comprobar teorías, sino en **estimar si un determinado tratamiento de naturaleza clínica, cognitiva, conductual, educativa, o de cualquier otra naturaleza, ejerce influencia sobre la(s) conducta(s) que se somete(n) a estudio**.

En este tipo de experimentos, las hipótesis estadísticas se formulan en los mismos términos que acabamos de describir. Sin embargo, el objetivo del investigador no consiste únicamente en rechazar la hipótesis de nulidad, sino en estimar también otra serie de estadísticos que puedan aportar información sobre la verdadera eficacia de los tratamientos.

Retomando el primero de los dos objetivos que persigue la experimentación, a saber, la *comprobación de teorías*, nos ha parecido interesante recoger una reflexión que realizan Pascual y Musitu (1981) acerca de tal objetivo. Partiendo del planteamiento clásico de Postman (1955) respecto a que «el sentido y la finalidad de la experimentación consiste en verificar las implicaciones de una proposición teórica», los autores afirman que la posibilidad de verificar los enunciados teóricos ha sido ampliamente refutada desde la filosofía de la ciencia. En este sentido, los análisis filosóficos ponen de manifiesto que, si lo entendemos como un procedimiento para justificar la validez de las teorías a partir de unos principios formales (Putnam, 1981), el método de verificar teorías no tiene justificación. No obstante, Pascual y Musitu consideran que la carencia de validez lógica no implica carencia de significado y función. En consecuencia, afirman que aunque el experimento no permite comprobar de forma objetiva los enunciados teóricos, posibilita su reconceptualización y, además, constituye una fuente de nuevas ideas e intuiciones. De esta forma, el experimento adquiere mayor relevancia en el contexto del «descubrimiento» que en el contexto de la «verificación». En nuestra opinión, ambos objetivos no son excluyentes sino complementarios. Por ello, creemos que la potencialidad del experimento para comprobar hipótesis o teorías, aunque sujeta siempre a cierto margen de error, es una de sus cualidades más relevantes y que no contradice en absoluto la capacidad que éste posee para generar nuevas ideas o predicciones.

3.2. CRITERIOS QUE GARANTIZAN LA CALIDAD METODOLÓGICA DEL EXPERIMENTO: EL PRINCIPIO MAX-MIN-CON

La calidad metodológica del experimento depende en gran medida de la planificación previa que haya hecho el investigador. Cuando planifica el estudio, el objetivo principal del investigador consiste en garantizar la precisión en el contraste de la hipótesis. Para ello, debe atender a una serie de criterios metodológicos propuestos por Kerlinger (1986), que se basan en la descomposición de la varianza total y que, en su conjunto, se conocen como *principio MAX-MIN-CON*. Antes de abordar las tres reglas fundamentales que configuran este principio, y que constituyen el principal objetivo metodológico de todo diseño eficaz, examinaremos el concepto de varianza y sus diferentes componentes.

El término **varianza total** hace referencia a la variabilidad que se observa en el registro de las puntuaciones de la variable dependiente bajo diferentes condiciones de tratamiento. La varianza total puede descomponerse en distintos tipos de varianza en función de su origen (véase la Figura 3.2). Cuando se lleva a cabo un experimento, su principal objetivo consiste en saber si la mayor parte de dicha varianza puede atribuirse a la manipulación de la(s) variable(s) independiente(s). El análisis de la varianza desde una perspectiva metodológica se centra en dilucidar tal cuestión.

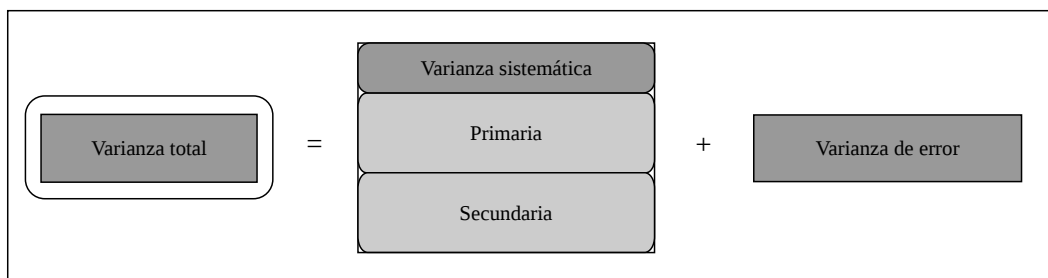


Figura 3.2 Descomposición de la varianza.

Kerlinger divide la *varianza total* de un conjunto de datos en dos amplias categorías: *varianza sistemática* y *varianza de error*.

La **varianza sistemática** hace referencia a la tendencia que presentan las puntuaciones de la variable dependiente a desviarse de su promedio en una determinada dirección. Como señala Arnau (1990b), esta variabilidad puede deberse a la influencia de la(s) variable(s) independiente(s) o a los efectos de factores extraños no controlados que son ajenos a los objetivos de la investigación. En función de estas dos posibilidades, la varianza sistemática se subdivide en otras dos fuentes de variación: la *varianza sistemática primaria* y la *varianza sistemática secundaria*.

La **varianza sistemática primaria**, conocida también como *varianza intergrupos*, *varianza experimental* o *varianza pretendida*, hace referencia a la variabilidad de la variable dependiente que refleja los efectos debidos a la(s) variable(s) independiente(s). La **varianza sistemática secundaria** o *varianza no pretendida*, por su parte, refleja la variabilidad de la variable dependiente asociada a la presencia de factores de perturbación o variables extrañas que no forman parte de la(s) hipótesis que se somete(n) a contrastación. Si alguna de tales variables extrañas varía junto con la(s) variable(s) independiente(s) sin que se detecte su presencia, los resultados del experimento quedan sesgados, siendo imposible determinar si

la varianza sistemática se debe a la influencia de la(s) variable(s) independiente(s) o a los efectos ejercidos por factores perturbadores ajenos a los objetivos de la investigación.

El segundo tipo de varianza, a saber, la **varianza de error** o *varianza intragrupo*, hace referencia al conjunto de fluctuaciones que, debido a factores aleatorios, presentan los datos en cualquier dirección con respecto a su promedio. Como señala Arnau (1990b), se trata de una varianza asistemática e impredecible que se estima en función de las diferencias existentes entre las puntuaciones dentro de cada grupo experimental. Las principales fuentes de varianza de error son las diferencias interindividuales, los errores en los sistemas de registro o de medida y los factores inherentes a la propia disposición experimental que afectan de forma diferencial a los sujetos.

Bajo este esquema, la calidad metodológica del experimento resulta óptima en la medida en que toda la varianza sistemática secundaria es extraída de la varianza total ya que, de esa forma, el cociente entre la varianza intergrupos y la varianza intragrupo refleja de manera adecuada la efectividad de los tratamientos experimentales. Por el contrario, el experimento no será eficaz si el investigador no puede controlar la varianza sistemática secundaria. En tal caso, dicha varianza pasa a formar parte de la varianza de error, y reduce en gran medida la *sensibilidad* de la investigación. Por tanto, en la planificación del estudio, el investigador debe seguir un principio metodológico de suma importancia, conocido como **principio MAX-MIN-CON**, que consta de tres reglas fundamentales:

1. MAXimización de la varianza sistemática primaria;
2. MINimización de la varianza de error;
3. CONtrol de la varianza sistemática secundaria.

El cumplimiento de este principio redundará en un aumento de la *sensibilidad* de la investigación que se deriva fundamentalmente de la maximización de la varianza sistemática primaria. A su vez, el control de la varianza sistemática secundaria y la minimización de la varianza de error garantizan que el experimento posea una adecuada *validez*.

La **maximización de la varianza sistemática primaria** consiste en potenciar al máximo la efectividad de los tratamientos. De acuerdo con la lógica de cálculo del análisis de la varianza, el efecto de un tratamiento se comprueba comparando el valor de la varianza intergrupos con el valor de la varianza intragrupo. En consecuencia, el aumento de la varianza intergrupos constituye un elemento decisivo para incrementar la potencia probatoria del diseño. En caso de que la relación funcional entre la variable independiente y la variable dependiente sea lineal, la maximización de dicha varianza se consigue mediante la selección de los valores extremos de la variable independiente. Para otro tipo de relaciones funcionales, tales como, por ejemplo, las relaciones monotónicas y curvilíneas, el mejor criterio para maximizar dicha varianza consiste en elegir valores óptimos de la variable manipulada (Arnau, 1990b; Ato, 1991).

La **minimización de la varianza de error** también permite incrementar la precisión en la inferencia de la hipótesis. Arnau (1990b) plantea dos procedimientos para reducir el tamaño del error. El primero de ellos consiste en controlar al máximo cualquier fuente de variación extraña mediante la elección de un adecuado diseño experimental, con lo que se obtiene una mayor precisión en la estimación de los efectos experimentales. En segundo lugar, el hecho de aumentar la homogeneidad entre las unidades de observación y la precisión en el registro de los datos también permiten minimizar la varianza de error. Con respecto a este último punto, el autor recomienda el uso de instrumentos de medida automáticos y estandarizados, dado que mediante este tipo de instrumentos es posible obtener registros con alto grado de precisión o fiabilidad.

Por último, el **control de la varianza sistemática secundaria** pretende evitar el *efecto de confundido* derivado de la influencia que pueden ejercer las variables extrañas sobre las respuestas de los sujetos. Como señala Arnau (1994), este objetivo se halla estrechamente vinculado al diseño experimental. En este sentido, existen modelos de diseño específicos orientados al control de la varianza sistemática secundaria o varianza no pretendida. Christensen (1988), Arnau (1990b), Ato (1991), Balluerka e Isasi (1996) y Balluerka (1999), entre otros, proporcionan una excelente síntesis de los procedimientos de control de variables extrañas utilizados habitualmente en el ámbito de la psicología.

4

PRINCIPALES CRITERIOS PARA LA CLASIFICACIÓN DE LOS DISEÑOS EXPERIMENTALES CLÁSICOS O FISHERIANOS

El **diseño experimental clásico**, conocido también como *diseño de $N > 1$* o *diseño de comparación de grupos* (Fisher, 1925, 1935), surgió al amparo de las ciencias sociales y de algunas ramas de las ciencias naturales, entre las que destaca la agricultura. Cook y Campbell (1986) utilizan el término *tradición del control estadístico* para conceptualizar este enfoque. Ello se debe a que se trata de un modelo de diseño con claras raíces estadísticas, en el que el sujeto individual en sí mismo carece de interés, siendo el grupo el principal elemento de referencia. Así, la comparación entre las puntuaciones medias de los distintos grupos se toma como base para extraer conclusiones acerca de la efectividad de los tratamientos experimentales. Debido a la importancia de partir de grupos de sujetos inicialmente equivalentes, la aleatorización constituye el principio fundamental del diseño fisheriano. Por otra parte, y como señala Arnau (1990b), el enfoque experimental clásico se fundamenta en el registro de una sola observación o de un conjunto reducido de observaciones por sujeto y enfatiza, casi con exclusividad, el resultado o el impacto final que produce el tratamiento sobre una determinada característica poblacional. A su vez, dado que toma como referente básico la teoría probabilística y que utiliza, en consecuencia, muestras de sujetos, posee como característica específica la posibilidad de generalizar los resultados a las poblaciones o universos de origen.

Aunque en la mayoría de los textos sobre el diseño experimental clásico se abordan los mismos modelos de diseño, existen múltiples categorizaciones al respecto. Estas clasificaciones parten de criterios muy diversos, pero no excluyentes. A continuación se exponen los criterios que se utilizan con mayor frecuencia en el ámbito de la psicología, así como los tipos de diseños que se distinguen en función de tales criterios.

A) Un criterio de extendido uso entre los investigadores experimentales hace referencia a la **estrategia empleada para la comparación entre los tratamientos administrados a los sujetos**. Atendiendo a este criterio, se pueden distinguir tres tipos de diseños:

- *Diseño intergrupos o intersujetos*. Cada tratamiento se administra a un grupo distinto de sujetos. De esta forma, los efectos de la(s) variable(s) independiente(s) se infieren comparando las medias obtenidas en la(s) variable(s) dependiente(s) por diferentes grupos de sujetos.
- *Diseño intrasujeto*. También se conoce con el nombre de *diseño de medidas repetidas* o *diseño de tratamientos $\times$ sujetos*. En este caso, cada uno de los tratamientos se administra al mismo grupo de sujetos. En consecuencia, la comparación entre los diversos niveles de tratamiento o entre las distintas condiciones experimentales se lleva a cabo dentro del mismo grupo de sujetos.
- *Diseño mixto o de medidas parcialmente repetidas*. Es un diseño que se caracteriza por la combinación de la estrategia intergrupos e intrasujeto. Por un lado, distintos grupos de sujetos se someten a diferentes tratamientos y sirven como referencia para comparar la efectividad de tales tratamientos. Por otro lado, cada uno de los grupos recibe toda una serie de tratamientos intrasujeto, de forma que la eficacia de este segundo conjunto de tratamientos se infiere realizando comparaciones dentro del mismo grupo. Debido a su naturaleza, los diseños mixtos pertenecen necesariamente a la categoría de diseños que se conocen como *diseños factoriales*, que abordaremos en el punto siguiente.

Los diseños intergrupos también se categorizan como *diseños de medida única*. A su vez, los diseños intrasujeto y los diseños mixtos suelen integrarse dentro de una categoría más amplia de diseños que se conocen como *diseños de medida múltiple*.

B) Otro criterio distintivo ampliamente utilizado en la investigación experimental se fundamenta en la **cantidad de variables independientes o factores de los que consta el diseño**. En función de este criterio, cabe distinguir dos grandes categorías de diseños:

- *Diseños simples o unifactoriales*. Son los diseños que constan de una sola variable independiente. En su estructura más sencilla, la variable independiente adopta sólo dos niveles, dando lugar a los diseños que, de forma genérica, se denominan *diseños bivalentes*. Si, además, la estructura es intergrupos, tales diseños se conocen como *diseños de dos grupos*. Por otra parte, los diseños unifactoriales en los que la variable independiente adopta más de dos niveles reciben el nombre de *diseños funcionales*, *multivalentes* o *multinivel* y, en caso de ser diseños intergrupos, se les asigna la denominación de *diseños multigrupos*.
- *Diseños complejos o factoriales*. Son aquellos diseños que constan de dos o más variables independientes o factores. Estos diseños se subdividen en distintas modalidades en función de criterios específicos, tales como la cantidad de valores adoptados por cada variable independiente o la configuración completa o incompleta de las combinaciones experimentales, así como en función de otros criterios taxonómicos comunes al resto de diseños experimentales.

C) Un tercer criterio de clasificación de gran utilidad en el contexto experimental hace referencia a la **técnica de control asociada a la estructura del diseño**. De acuerdo con este criterio, el diseño experimental se subdivide en tres amplios bloques, cada uno de los cuales incluye varios tipos de diseños. Tales bloques son:

- *Diseños de grupos completamente aleatorios*. Son aquellos diseños en los que se utiliza la técnica de control denominada *aleatorización*. En otras palabras, se trata de estructuras de diseño en las que los sujetos se asignan de forma completamente aleatoria a los distintos niveles de tratamiento.

- *Diseños con una o más dimensiones de bloqueo, diseños emparejados y diseños jerárquicos.* En esta categoría se incluyen los modelos de diseño asociados con distintos procedimientos de *equilibración*. Aunque la lógica subyacente a todos ellos es la misma, el control se lleva a cabo de forma específica en cada modalidad de diseño. Así, aquellos diseños en los que se aplica la *técnica de bloqueo* con el objetivo de controlar la influencia de una variable extraña, se conocen como *diseños de bloques aleatorios*. En caso de utilizarse el *doble bloqueo* o el *triple bloqueo*, los diseños se denominan, respectivamente, *diseños de cuadrado latino* y *diseños de cuadrado grecolatino*. El uso de la *técnica de emparejamiento* como procedimiento de control se asocia a los diseños que reciben el nombre de *diseños emparejados*. Por último, los *diseños jerárquicos* son aquellos diseños vinculados al método de control experimental que se conoce como *técnica de anidamiento*.
- *Diseños intrasujeto.* Estos diseños ya han sido descritos al exponer el criterio referido a la *estrategia empleada para la comparación entre los tratamientos administrados a los sujetos*. No obstante, se incluyen también bajo el criterio que aquí nos ocupa porque en tales diseños, el propio sujeto se convierte en instrumento de bloqueo o de control para reducir todo tipo de varianza no asociada con la manipulación experimental.

D) Un cuarto criterio taxonómico que, debido fundamentalmente a la extensión en el uso de los paquetes estadísticos para el análisis de datos por ordenador, posee hoy en día una relevancia especial, es el asociado a la **cantidad de variables dependientes incluidas en el diseño**. En función de este criterio, cabe distinguir entre *diseños univariantes* o *univariados* y *diseños multivariantes* o *multivariados*. Los *primeros* son aquellos diseños en los que se registra una sola variable dependiente y poseen las características propias de los diseños experimentales clásicos o fisherianos tal y como éstos se propusieron en sus inicios. Los *diseños multivariantes*, por su parte, se caracterizan por registrar varias variables dependientes simultáneamente.

E) La **configuración completa o incompleta de las combinaciones experimentales** es otro de los criterios que nos permite distinguir entre diferentes modalidades de diseño. Los *diseños completos* son aquellos que poseen unidades experimentales en todas las combinaciones de tratamientos. Los *diseños incompletos*, por el contrario, son las estructuras caracterizadas por no disponer de unidades experimentales en alguna combinación de tratamientos. En función del motivo por el que se carece de dichas combinaciones, es posible subdividir los diseños incompletos en dos categorías: *diseños estructuralmente incompletos* y *diseños accidentalmente incompletos*. En los *primeros*, la omisión de ciertas combinaciones tiene una justificación de naturaleza metodológica. En los *segundos*, por el contrario, se carece de tales combinaciones porque, simplemente, no se han podido administrar, lo que posee claras implicaciones de cara al análisis estadístico de los datos.

Dentro de los diseños estructuralmente incompletos, cabe destacar cinco modelos de diseño: los diseños jerárquicos, los diseños de cuadrado latino y de cuadrado grecolatino intersujetos, los diseños de bloques incompletos y los diseños fraccionados. El diseño jerárquico, o con variables anidadas, es una modalidad de diseño en la que los niveles de un factor, que normalmente es de clasificación y que se denomina factor anidado, están representados en un solo nivel de otro factor de naturaleza experimental. El diseño de cuadrado latino intersujetos se caracteriza por utilizar dos factores de clasificación con el objetivo de controlar dos variables extrañas mediante la técnica de bloqueo. En este diseño, los niveles de un factor experimental se administran en cada una de las categorías de dos factores de bloqueo, por lo que resulta necesario que todos los factores posean el mismo número de niveles. El

diseño de cuadrado grecolatino sigue la misma lógica que el anterior, con la única diferencia de que en este último se utilizan tres variables de bloqueo. El diseño de bloques incompletos es un tipo de diseño factorial en el que se omite la estimación de determinados efectos factoriales, normalmente interacciones de escasa importancia, con el objetivo de eliminar el influjo de alguna variable extraña. Por último, el diseño fraccionado es aquel diseño en el que el número de combinaciones experimentales se reduce a una fracción de la cantidad total de tales combinaciones.

F) Centrándonos en criterios taxonómicos estrechamente vinculados con los análisis estadísticos cabe citar, en primer lugar, el referido al **tipo de técnica utilizada para analizar los datos**. En función de dicho criterio, los diseños suelen dividirse en *diseños paramétricos* y *diseños no paramétricos*. Los *primeros* son aquellos diseños que utilizan técnicas de análisis basadas en una serie de estrictas suposiciones acerca de la naturaleza de la población de la que se obtienen los datos. Los *diseños no paramétricos*, por el contrario, se vinculan a las técnicas conocidas como *técnicas no paramétricas* o *de distribución libre*, en las que los supuestos sobre los parámetros poblacionales son mucho menos severos. En opinión de Arnau (1986), mediante esta distinción de carácter meramente estadístico, el concepto de diseño no se considera en toda su amplitud, quedando restringido o limitado a tan sólo uno de los elementos que lo configuran.

G) Tomando como criterio de clasificación las **posibilidades de control estadístico que brinda el diseño**, cabe distinguir entre el *diseño con covariables* o *diseño de covarianza* y el *diseño sin covariables*. El *diseño con covariables* es una modalidad de diseño caracterizada por controlar una o más variables perturbadoras mediante la técnica de ajuste estadístico que se conoce como *análisis de la covarianza*. El *diseño sin covariables*, por el contrario, es aquel en el que no se aplica dicho procedimiento de control sobre ninguna variable extraña.

H) Por último, citaremos el criterio que hace referencia a la **constancia en la cantidad de observaciones por combinación de tratamientos**. Este criterio permite dividir los diseños en dos grandes categorías: *diseños equilibrados* y *diseños no equilibrados*¹. Los *primeros* son aquellos diseños que tienen la misma cantidad de sujetos en todas las combinaciones experimentales. En los *diseños no equilibrados*, por el contrario, el número de sujetos no se mantiene constante a lo largo de todas las combinaciones de tratamientos.

En la Tabla 4.1 se presentan los criterios de clasificación del diseño experimental clásico previamente expuesto.

¹ Aunque en el presente texto se desarrollan análisis de la varianza para diseños equilibrados, los lectores interesados en profundizar en los problemas que plantea la elección del método para calcular las sumas de cuadrados del ANOVA en los diseños no equilibrados y, de forma más genérica, en los análisis asociados a tales diseños, pueden consultar el reciente artículo de Macnaughton (1998) y la excelente obra de Searle (1987).

TABLA 4.1 Criterios para la clasificación del diseño experimental clásico o fisheriano (diseño de $N > 1$)

Criterio	Estrategia	Diseño	Subclasificaciones
A) Estrategia de comparación entre tratamientos.	Cada tratamiento se administra a un grupo distinto de sujetos.	Intergrupos o intersujetos o de medida única.	
	Cada tratamiento se administra al mismo grupo de sujetos.	Intrasujeto o de medidas repetidas o tratamientos $\times$ sujetos.	
	Combinación de las estrategias intergrupos e intrasujeto.	Mixto o de medidas parcialmente repetidas.	
	Una sola variable independiente.	Simples o unifactoriales.	• VI con dos niveles; D. bivalentes (en caso de ser intergrupos se denominan D. de dos grupos). • VI con más de dos niveles; D. funcionales, multivalentes o multinivel (en caso de ser intergrupos se denominan D. multigrupos).
C) Técnica de control asociada a la estructura del diseño.	Dos o más variables independientes.	Complejos o factoriales.	En función de la cantidad de valores adoptados por cada VI y de otros criterios taxonómicos.
	Técnica de control: aleatorización.	De grupos completamente aleatorios.	
	Técnica de control: equilibración.	Diseños con una o más dimensiones de bloque. Diseños emparejados. Diseños jerárquicos.	
	Técnica de bloqueo simple. Técnica de doble bloqueo. Técnica de triple bloqueo. Técnica de emparejamiento. Técnica de anidamiento.	Diseños de bloques aleatorios. Diseños de cuadrado latino. Diseños de cuadrado grecolatino. Diseños emparejados. Diseños jerárquicos.	
D) Cantidad de variables dependientes.	Técnica de control: el propio sujeto.	Diseños intrasujeto.	
	Una variable dependiente.	Univariantes o univariados.	
	Varias variables dependientes.	Multivariantes o multivariados.	

TABLA 4.1 Criterios para la clasificación del diseño experimental clásico o fisheriano (diseño de $N > 1$) (continuación)

Criterio	Estrategia	Diseño	Subclasificaciones
E) Configuración completa o incompleta de las combinaciones experimentales.	Presencia de unidades experimentales en todas las combinaciones de tratamientos.	Completo.	
	Ausencia de unidades experimentales en una o varias combinaciones de tratamientos.	Incompleto.	• Ausencia por imposibilidad de administrar tratamientos; <i>D. accidentalmente incompletos.</i> • Ausencia con justificación metodológica: <i>D. estructuralmente incompletos:</i> * Los niveles de un factor (anidado) están representados en un solo nivel de otro factor (experimental); <i>D. jerárquico o con variables anidadas.</i> * Inclusión de dos variables de bloqueo: <i>D. de cuadrado latino intersujetos.</i> * Inclusión de tres variables de bloqueo: <i>D. de cuadrado grecolatino.</i> * Omisión de determinados efectos factoriales (control); <i>D. de bloques incompletos.</i> * Reducción del número de combinaciones factoriales a una fracción de la cantidad total; <i>D. fraccionado.</i>
F) Tipo de técnica utilizada para el análisis de datos.	Basada en estrictos supuestos acerca de los parámetros poblacionales.	Paramétrico.	
	Supuestos menos severos sobre los parámetros poblacionales.	No paramétrico o de distribución libre.	
G) Control estadístico.	Mediante ajuste estadístico de variables perturbadoras.	Con covariables o de covarianza.	
	No utilización de ajuste estadístico.	Sin covariables.	
H) Constancia en el número de observaciones.	Mismo n° de sujetos por tratamiento.	Equilibrado.	
	Diferente n° de sujetos por tratamiento.	No equilibrado.	

5

APROXIMACIÓN GENERAL A LA TÉCNICA DE ANÁLISIS DE DATOS MÁS UTILIZADA EN EL ÁMBITO DEL DISEÑO EXPERIMENTAL CLÁSICO: EL ANÁLISIS DE LA VARIANZA

En los capítulos precedentes hemos intentado definir los aspectos esenciales del diseño experimental y ubicarlo en el marco de las metodologías que se utilizan habitualmente dentro de las ciencias del comportamiento. Para finalizar con esta primera aproximación al diseño experimental clásico o fisheriano, nos ha parecido interesante describir, de forma genérica, los aspectos más relevantes de la técnica estadística que se aplica por excelencia cuando se analizan los datos procedentes de este tipo de diseño, a saber, del análisis de la varianza. Por ello, en el presente apartado abordamos los supuestos, los objetivos, las fases y otra serie de aspectos que nos ayudan a comprender la lógica que subyace a dicha estrategia analítica. No obstante, queremos dejar claro que lo que aquí se le ofrece al lector no es sino una aproximación general al análisis de la varianza, cuya finalidad radica en mostrar, a grandes rasgos, en qué consiste dicha técnica. En epígrafes posteriores aplicamos distintos modelos de análisis de la varianza a los datos derivados de las modalidades de diseño que presentamos, a fin de que el lector disponga de ejemplos o de aplicaciones concretas de esta estrategia en diferentes situaciones de investigación.

5.1. ANÁLISIS UNIVARIADO DE LA VARIANZA: ANOVA

El análisis de la varianza ocupa un lugar privilegiado entre las técnicas estadísticas asociadas al diseño experimental clásico. Independientemente de la cantidad de factores de los que conste el diseño, la prueba de la hipótesis se puede realizar con modelos que incluyen más de una variable dependiente o una sola variable dependiente. En este último caso, la prueba estadística se denomina *análisis univariado de la varianza* y su objetivo consiste en comprobar si uno o más tratamientos experimentales producen un efecto determinado sobre la variable dependiente. Esta potencial relación causal entre la(s) variable(s) independiente(s)

y la variable dependiente puede describirse mediante una ecuación matemática de carácter lineal que se conoce como *modelo estructural matemático del ANOVA*. Siguiendo a Arnau (1986) y a Namboodiri, Carter y Blalock (1975), la estructura básica que subyace al análisis de la varianza puede expresarse de la siguiente forma (véase la Figura 5.1).

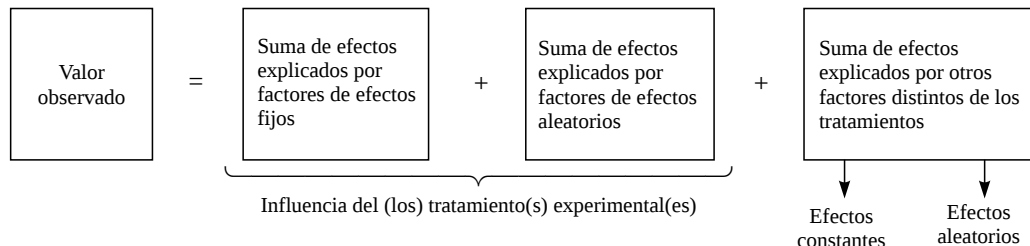


Figura 5.1 Estructura básica que subyace al análisis de la varianza.

Según este modelo, toda observación experimental puede ser explicada como la suma de una serie de efectos de factores causales. Los dos primeros términos del miembro derecho de la igualdad representan el efecto de un conjunto de factores o variables que pueden tener un número fijo de valores o, por el contrario, pueden ser de naturaleza aleatoria. El tercer término recoge el efecto combinado de todos aquellos factores diferentes del tratamiento. Algunos de tales factores permanecen constantes a lo largo del experimento y se expresan mediante el símbolo μ . Otros, por el contrario, fluctúan de manera totalmente aleatoria, llegando a ser impredecibles a partir de los factores. Este conjunto de variables incontrolables se designa mediante el símbolo ε_{ij} y constituye el error experimental.

Siguiendo el esquema anterior, la estructura básica del modelo del análisis de la varianza, en su expresión más simple, viene dada por la siguiente ecuación:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (5.1)$$

donde:

- y_{ij} = Puntuación obtenida en la variable dependiente por el sujeto i en el tratamiento j .
- μ = Puntuación media de la población de la que se ha extraído la observación. Este término es el equivalente a β_0 en los modelos de regresión y constituye un elemento constante para todos los datos del experimento.
- α_j = Efecto del j -ésimo tratamiento experimental. Si el modelo estructural es de efectos fijos, este componente es constante para todas las observaciones obtenidas dentro de cada grupo de tratamiento.
- ε_{ij} = Error correspondiente a la i -ésima observación bajo el j -ésimo tratamiento. Recoge la fluctuación entre las diferentes observaciones realizadas bajo la acción de un mismo tratamiento.

Como todo modelo matemático, el que se describe en esta ecuación es totalmente teórico y sus parámetros deben derivarse de las observaciones concretas de cada uno de los sujetos del experimento. El procedimiento matemático utilizado habitualmente en el ámbito de la psicología, para la estimación de los parámetros del modelo, se conoce como el *criterio de los mínimos cuadrados ordinarios (MCO)* y consiste en minimizar la suma de cuadrados del componente de error del modelo. En definitiva, la finalidad del análisis de la varianza radica en comprobar si alguno de los parámetros del modelo se desvía suficientemente de cero y,

por ende, si alguno de los tratamientos experimentales ejerce una influencia significativa sobre el fenómeno objeto de estudio. Para realizar tal comprobación se siguen una serie de etapas que conducen a la estimación de la razón F , razón en la que se fundamenta la prueba de la hipótesis. A continuación describimos de forma sucinta tales etapas así como los elementos que intervienen en cada una de ellas. Tomamos como referencia un modelo simple, a saber, el diseño unifactorial multigrupos al azar. Este modelo consta de una sola variable independiente que adopta k valores o niveles de tratamiento. El esquema general del diseño se representa en la Tabla 5.1.

TABLA 5.1 Esquema general del diseño unifactorial multigrupos al azar

		Variable independiente						
		A						
		A_1	A_2	$\cdots$	A_j	$\cdots$	A_k	$\bar{y}_{.i}$
Sujetos	S_1	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	y_{1k}	$\bar{y}_{1.}$
	S_2	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	y_{2k}	$\bar{y}_{2.}$
	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
	S_i	y_{i1}	y_{i2}	$\cdots$	y_{ij}	$\cdots$	y_{ik}	$\bar{y}_{i.}$
	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
	S_n	y_{n1}	y_{n2}	$\cdots$	y_{nj}	$\cdots$	y_{nk}	$\bar{y}_{n.}$
$\bar{y}_{.j}$		$\bar{y}_{.1}$	$\bar{y}_{.2}$	$\cdots$	$\bar{y}_{.j}$	$\cdots$	$\bar{y}_{.k}$	$\bar{y}_{..}$

Como se observa en la Tabla 5.1, el diseño consta de n sujetos u observaciones en cada nivel de tratamiento. La puntuación obtenida por cada sujeto en la variable dependiente se representa mediante la letra «y» acompañada de dos subíndices. El primero hace referencia a los distintos sujetos dentro de cada grupo de tratamiento ($i = 1, 2, \dots, n$) y el segundo especifica el tratamiento que le corresponde a cada sujeto experimental ($j = 1, 2, \dots, k$).

La ecuación estructural asociada a este diseño es la que hemos visto anteriormente (5.1). Por su parte, la aplicación del procedimiento de los *mínimos cuadrados ordinarios* para la estimación de los parámetros requiere que se hagan las siguientes suposiciones acerca del componente de error del modelo:

1. Los ε_{ij} deben distribuirse normalmente.
2. Los ε_{ij} han de ser independientes entre sí.
3. La media de los ε_{ij} debe ser igual a cero.
4. La varianza de los ε_{ij} ha de coincidir con la varianza de la población:

$$V(\varepsilon_{ij}) = \sigma_\varepsilon^2$$

Si se aplica el criterio de mínimos cuadrados a partir de estos supuestos, los parámetros pueden estimarse de la siguiente forma:

$$\begin{aligned} \bar{y}_{..} &\rightarrow \mu \\ \bar{y}_{.j} - \bar{y}_{..} &= \mu_j - \mu \rightarrow \alpha_j \end{aligned} \quad (5.2)$$

Sustituyendo dichos valores en la ecuación estructural del modelo (Ecuación 5.1):

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{.j} - \bar{y}_{..}) + \varepsilon_{ij} \quad (5.3)$$

En consecuencia,

$$\varepsilon_{ij} = y_{ij} - \bar{y}_{.j} \quad (5.4)$$

Introduciendo esta última equivalencia en el modelo:

$$y_{ij} = \bar{y}_{..} + (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{.j}) \quad (5.5)$$

Colocando y_{ij} en el miembro izquierdo de la igualdad, la desviación de una observación en torno a la media de la población se expresa como:

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{.j}) \quad (5.6)$$

Esta ecuación constituye el *principio básico del análisis de la varianza*. De acuerdo con dicho principio, la distancia (o tamaño de variación) de un valor observado concreto, y_{ij} , con respecto a la media de todos los valores observados, $\bar{y}_{..}$, se puede dividir en dos componentes:

- a) La amplitud de la variación de la media del grupo de tratamiento j con respecto a la media global, $\bar{y}_{.j} - \bar{y}_{..}$, la cual representa la desviación causada por el tratamiento experimental o la *desviación explicada*.
- b) La amplitud de la variación de cada observación con respecto a la media de su grupo de tratamiento, $y_{ij} - \bar{y}_{.j}$, que representa la desviación debida a factores extraños no controlados por el experimentador o la *desviación no explicada*.

El propósito básico del ANOVA radica en calcular estos dos componentes de variación y en convertirlos en valores comparables. Para ello, se han de estimar en primer lugar las denominadas *sumas de cuadrados* o variaciones que se producen en un experimento. Veamos cómo se realiza dicha estimación.

5.1.1. Sumas de cuadrados

Las sumas de cuadrados reflejan las diferentes variaciones que tienen lugar en un experimento y constituyen un elemento fundamental para el cálculo de las varianzas.

Partiendo de una sola observación, y_{ij} , y aplicando la descomposición de la *desviación total* (puntuaciones diferenciales con respecto a la media) en los componentes que acabamos de describir como *desviación explicada* y *desviación no explicada*, tenemos:

$$y_{ij} - \bar{y}_{..} = (\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{.j}) \quad (5.6)$$

Elevando al cuadrado ambos miembros de la igualdad:

$$(y_{ij} - \bar{y}_{..})^2 = [(\bar{y}_{.j} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{.j})]^2 \quad (5.7)$$

Desarrollando el binomio al cuadrado del segundo miembro de la igualdad:

$$(y_{ij} - \bar{y}_{..})^2 = (\bar{y}_{.j} - \bar{y}_{..})^2 + (y_{ij} - \bar{y}_{.j})^2 + 2(\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{.j}) \quad (5.8)$$

Generalizando la ecuación para todo i y j del experimento:

$$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = n \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_i \sum_j (y_{ij} - \bar{y}_{.j})^2 + 2 \sum_i \sum_j (\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{.j}) \quad (5.9)$$

Dado que la suma de las variaciones de los y_{ij} en torno a su media, $\bar{y}_{..}$, es igual a cero:

$$2 \sum_i \sum_j (\bar{y}_{.j} - \bar{y}_{..})(y_{ij} - \bar{y}_{.j}) = 0 \quad (5.10)$$

En consecuencia, se obtiene la siguiente igualdad:

$$\sum_i \sum_j (y_{ij} - \bar{y}_{..})^2 = n \sum_j (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_i \sum_j (y_{ij} - \bar{y}_{.j})^2 \quad (5.11)$$

Los términos de esta ecuación corresponden a lo que se conoce, respectivamente, como **suma cuadrática total**, **suma cuadrática intergrupos** y **suma cuadrática intragrupo**. Debido a que el cálculo de las sumas de cuadrados mediante las fórmulas anteriores resulta complejo, dicha estimación suele llevarse a cabo aplicando una serie de fórmulas equivalentes, que para el diseño unifactorial multigrupos aleatorios son:

Suma cuadrática total:

$$\sum_i \sum_j y_{ij}^2 - \frac{\left(\sum_i \sum_j y_{ij} \right)^2}{kn} \quad (5.12)$$

Suma cuadrática intergrupos, intertratamientos o explicada:

$$\sum_j \frac{\left(\sum_i y_{ij} \right)^2}{n} - \frac{\left(\sum_i \sum_j y_{ij} \right)^2}{kn} \quad (5.13)$$

Suma cuadrática intragrupo, residual o no explicada:

$$\sum_i \sum_j y_{ij}^2 - \sum_j \frac{\left(\sum_i y_{ij} \right)^2}{n} \quad (5.14)$$

Como se acaba de señalar, el cálculo de cada una de las sumas de cuadrados en función de las fórmulas anteriores sólo puede aplicarse al diseño multigrupos al azar con una variable independiente. En caso de que se emplee otro tipo de diseño, tanto la descomposición de la variación total como el cálculo de las sumas de cuadrados, debe realizarse partiendo de la ecuación estructural correspondiente al diseño que se utiliza como esquema de trabajo.

En términos generales, la suma de cuadrados intragrupo refleja la variabilidad de las puntuaciones de los sujetos dentro de un mismo nivel de tratamiento. Dado que en cada grupo de tratamiento los sujetos se someten a las mismas circunstancias, se supone que esta variabilidad es de naturaleza aleatoria. La suma cuadrática intergrupos, por su parte, hace referencia a la variabilidad producida por la acción de los diferentes tratamientos que reciben los sujetos a lo largo del experimento.

Una vez obtenidas la suma cuadrática total, intergrupos e intragrupo, han de compararse tales variaciones entre sí, para lo cual resulta necesario obtener sus correspondientes varianzas. La varianza es una variación promedio y, más concretamente, se define como la cantidad de variación por grado de libertad. En consecuencia, para proceder al cálculo de las varianzas se deben extraer, en primer lugar, los grados de libertad correspondientes a cada suma de cuadrados.

5.1.2. Grados de libertad

Arnau (1986) define los grados de libertad como la cantidad de observaciones básicas independientes de las que se dispone en la estimación de una fuente de variación. Así, siempre que se aplique una restricción a las observaciones básicas se pierde un grado de libertad. Si tales observaciones son los datos del experimento y éstos se utilizan para calcular la suma cuadrática del error, en la estimación de esta fuente de variación se impone una restricción dentro de cada uno de los grupos de tratamiento, ya que los datos deben variar en torno a la media del grupo. Dicha restricción hace que se pierda un grado de libertad por grupo. En consecuencia, los **grados de libertad de la fuente de error** se calculan restando una unidad a la suma de las observaciones registradas en cada nivel de tratamiento y multiplicando dicho valor por la cantidad de niveles que adopta la variable independiente, a saber:

$$gl_{\text{error}} = k(n - 1) = N - k \quad (5.15)$$

Para calcular los grados de libertad correspondientes a la suma cuadrática intergrupos se aplica el mismo razonamiento. Así, en este caso, la pérdida de un grado de libertad se debe a la restricción impuesta por la media global del experimento. En concreto, los **grados de libertad de la fuente de variación intergrupos** se calculan restando una unidad al número de niveles de la variable independiente, a saber:

$$gl_{\text{entre}} = k - 1 \quad (5.16)$$

Finalmente, los **grados de libertad correspondientes a la suma cuadrática total** se estiman restando una unidad al número total de observaciones de la variable dependiente. Como cabe observar, junto a la descomposición que se produce en la variación o suma de cuadrados total, los grados de libertad totales también se descomponen en dos términos: los grados de libertad correspondientes a la suma de cuadrados intergrupos y los asociados a la suma cuadrática del error:

$$\begin{aligned} gl_{\text{total}} &= gl_{\text{entre}} + gl_{\text{error}} \\ N - 1 &= (k - 1) + k(n - 1) \end{aligned} \quad (5.17)$$

Tras la obtención de los grados de libertad, se procede al cálculo de las varianzas o medias cuadráticas.

5.1.3. Varianzas o medias cuadráticas

Las varianzas o medias cuadráticas se obtienen dividiendo las variaciones o sumas de cuadrados entre sus correspondientes grados de libertad. De esta forma, se estima la proporción de varianza que corresponde a cada una de las fuentes de variación del modelo.

En el análisis univariado de la varianza hay que calcular dos medias cuadráticas. La primera corresponde a la fuente de variación intergrupos y se obtiene a partir de la variabilidad de las medias de los grupos de tratamiento. Su estimación se realiza aplicando la siguiente fórmula:

$$MC_{\text{entre}} = \frac{SC_{\text{entre}}}{gl_{\text{entre}}} = \frac{SC_{\text{entre}}}{k - 1} \quad (5.18)$$

No obstante, la varianza intergrupos no proporciona información suficiente para llevar a cabo la prueba de la hipótesis de nulidad, ya que dicha prueba también requiere obtener una estimación de las fluctuaciones debidas al azar, independientemente de la acción de los tratamientos. Por ello, se debe calcular la varianza de los datos dentro de cada grupo de tratamiento, es decir, la media cuadrática intragrupo o del error. Dicha varianza se obtiene aplicando la siguiente fórmula:

$$MC_{\text{error}} = \frac{SC_{\text{error}}}{gl_{\text{error}}} = \frac{SC_{\text{error}}}{k(n - 1)} \quad (5.19)$$

Dado que las varianzas están promediadas en función de los grados de libertad, resulta posible compararlas entre sí y esta comparación permite estimar la razón F .

5.1.4. Razón F

La razón F es el cociente entre la media cuadrática intergrupos o varianza explicada por el tratamiento y la media cuadrática del error o varianza residual, a saber:

$$F = \frac{MC_{\text{entre}}}{MC_{\text{error}}} \quad (5.20)$$

Bajo el supuesto de la *hipótesis nula* (H_0) o de la ausencia de diferencia entre los diversos grupos de tratamiento, la distribución muestral del estadístico se aproxima a la distribución F con $k - 1$ grados de libertad en el numerador y $k(n - 1)$ grados de libertad en el denominador.

Por tanto, este estadístico puede utilizarse para contrastar la hipótesis de igualdad entre las diferentes medias de la población, es decir, para saber si las medias de los grupos han sido extraídas de una misma población o de poblaciones idénticas. Si las medias de los grupos no difieren entre sí, la razón se acerca a la unidad, variando únicamente en función de las fluctuaciones del muestreo. Si, por el contrario, las medias de los grupos difieren entre sí, la razón excede la unidad en mayor medida de lo que podría atribuirse al mero azar. El rechazo de la H_0 significa que una de las medias presenta una gran diferencia con respecto a las demás, o que todas las medias son muy diferentes entre sí. En este caso, la media cuadrática intergrupos es significativamente mayor que la media cuadrática del error, lo que explica que el valor de F sea superior a la unidad. La H_0 se rechaza cuando el valor de F obtenido por el investigador (F empírica u observada) supera el valor crítico o tabular del estadístico F (F teórica) asumiendo un determinado margen de error de tipo I (α).

Cuando el modelo es de efectos fijos, los datos se pueden analizar aplicando un método alternativo que permite apreciar, de forma conjunta, el fundamento del modelo de regresión

y del modelo de diseño experimental. Se trata del **análisis de la varianza aplicado a la regresión**, cuyo objetivo consiste, al igual que el del análisis de la varianza, en descomponer la *variación total* en dos elementos:

- a) la *variación explicada* por el modelo de regresión, la cual se calcula a través del coeficiente de regresión estimado β' ;
- b) la *variación no explicada* por la regresión o variación residual.

A partir de esta descomposición de la variación total, similar a la que se realiza en el análisis de la varianza, y aplicando dicha técnica se dispone de una prueba estadística para validar el modelo: se trata de verificar si la *variación explicada* es superior a la *variación no explicada*. Arnau (1986) y Domenech y Riba (1985) exponen la lógica de este procedimiento y presentan diversos ejemplos en los que se calculan las sumas de cuadrados, las medias cuadráticas y la razón entre varianzas a partir del modelo de regresión.

Por último, es importante señalar que la única información que proporciona el rechazo de la hipótesis de nulidad en el análisis de la varianza es que las diferencias entre las medias de los tratamientos son mayores de lo que cabe esperar por azar. No obstante, en el caso de diseños con más de dos tratamientos, no permite saber entre qué medias específicas existen diferencias estadísticamente significativas. Por ello, cuando se rechaza la hipótesis nula deben someterse a prueba una serie de hipótesis específicas para interpretar adecuadamente los resultados. Estas hipótesis reciben el nombre de **contrastes de medias o comparaciones múltiples**.

Las comparaciones múltiples entre las medias de los tratamientos pueden ser de dos tipos, a saber: *comparaciones planificadas o a priori* y *comparaciones no planificadas o a posteriori*. A su vez, las comparaciones planificadas se subdividen en dos categorías: *contrastes no ortogonales* y *contrastes ortogonales o independientes entre sí*. La principal desventaja de las *comparaciones a priori* radica en que, a medida que aumenta el número de contrastes, también se incrementa la probabilidad de cometer un error de tipo I, es decir, la probabilidad de rechazar la H_0 siendo verdadera. Aunque existen diversos métodos (por ejemplo, la *corrección de Bonferroni*) para solventar este problema, el procedimiento más apropiado para evitar la tasa de error en las comparaciones múltiples consiste en utilizar *comparaciones no planificadas o a posteriori*, la mayor parte de las cuales fija la tasa de error al mismo nivel de significación para todos los pares de comparaciones (Klockars y Sax, 1985). A diferencia de las comparaciones que se plantean antes de llevar a cabo el análisis de la varianza, las comparaciones no planificadas se formulan en función de los resultados obtenidos en dicho análisis y, entre ellas, las más utilizadas son la *prueba HSD de Tukey*, el *procedimiento de Dunnett* y la *prueba de Scheffé*. Las principales técnicas de comparación múltiple pueden consultarse en Hochberg y Tamhane (1987), Keppel (1982), Toothaken (1991) y Winer, Brown y Michels (1991). En el Capítulo 6 se abordarán de forma más exhaustiva los contrastes de medias utilizados habitualmente en el ámbito de la psicología.

5.2. ANÁLISIS MULTIVARIADO DE LA VARIANZA: MANOVA

5.2.1. Relaciones entre el ANOVA y el MANOVA

El análisis multivariado de la varianza (MANOVA) es una generalización del análisis univariado de la varianza (ANOVA). La principal distinción entre ambos radica en que mientras el ANOVA se centra en el estudio de las diferencias entre las medias poblacionales de

los distintos grupos en una sola variable dependiente, el MANOVA examina tales diferencias en dos o más variables dependientes simultáneamente. No obstante, cabe señalar que, al igual que el ANOVA, el MANOVA sólo proporciona información respecto a la existencia de diferencias estadísticamente significativas entre las medias de los grupos en el conjunto de las variables dependientes, requiriendo la realización de pruebas complementarias para interpretar de una forma más precisa los resultados.

Bray y Maxwell (1993) destacan varias razones que, a su juicio, hacen aconsejable la utilización del MANOVA cuando se examinan diferencias entre medias. En primer lugar, los autores plantean que la mayoría de los investigadores no están interesados en evaluar las diferencias entre las medias de las puntuaciones correspondientes a una sola variable dependiente, sino a un conjunto de variables dependientes. Ello se debe a que normalmente son varios los constructos o comportamientos que pueden estar bajo el influjo de los tratamientos y, por tanto, la medición de más de una variable dependiente proporciona información más fiable respecto a los verdaderos efectos de la(s) variable(s) manipulada(s) por el investigador.

Por otra parte, si se realizan múltiples ANOVAs, se presupone que no existe correlación entre las variables dependientes o que dicha correlación no reviste interés. El MANOVA, por el contrario, permite examinar las diferencias entre las medias de los tratamientos en todas las variables dependientes conjuntamente, teniendo en cuenta la correlación existente entre ellas. Si dicha correlación es significativa, la realización de pruebas de hipótesis que incluyen una única variable dependiente lleva a un incremento en la tasa de error de tipo I. Además, la protección contra el error de tipo I que proporciona el MANOVA es mayor a medida que aumenta el grado de correlación que existe entre las variables dependientes. Por otra parte, cuando existe correlación entre tales variables, el MANOVA tiene una potencia superior al ANOVA. Tabachnick y Fidell (1989) añaden a las anteriores ventajas del MANOVA una adicional, a saber, que bajo ciertas condiciones, el análisis multivariado de la varianza permite apreciar diferencias que no pueden detectarse realizando análisis separados con cada una de las variables dependientes incluidas en el diseño.

Los argumentos que se acaban de esgrimir a favor del MANOVA no deben llevarnos a creer que esta técnica de análisis resulta siempre más adecuada que el ANOVA. De hecho, el análisis univariado de la varianza posee mayor potencia que el multivariado cuando no existe correlación entre las variables dependientes y cuando el tamaño de la muestra empleada en la investigación es pequeño. Por otra parte, en relación con la estimación del tamaño del efecto que se obtiene a partir de varias pruebas univariadas o de una sola prueba multivariada, Pascual, García y Frías (1995) consideran que no puede afirmarse categóricamente que la existencia de correlación entre las variables dependientes mejore la estimación del tamaño del efecto en el diseño multivariado. Así, en opinión de los autores, la diferencia en la estimación que puede obtenerse a partir de ambas estrategias de análisis está mediada por tres parámetros: la proporción de varianza que explica cada una de las variables dependientes por separado, la correlación entre estas variables y la correlación entre los residuales intragrupo.

5.2.2. Prueba de la hipótesis de nulidad en los diseños multivariados

A diferencia del ANOVA, el MANOVA posee más de un índice o estadístico para contrastar la hipótesis de nulidad multivariante. Tales índices se analizan al final del apartado, pero antes de proceder a su estudio, abordamos de forma genérica la lógica y los principales elementos de la prueba de significación multivariante.

La diferencia entre ambas matrices produce como resultado la matriz correspondiente a la *suma de cuadrados intergrupos (B)*:

$$B = \begin{bmatrix} \sum_j n_j (\bar{y}_{.j1} - \bar{y}_{..1})^2 & \sum_j n_j (\bar{y}_{.j1} - \bar{y}_{..1})(\bar{y}_{.j2} - \bar{y}_{..2}) \\ \sum_j n_j (\bar{y}_{.j1} - \bar{y}_{..1})(\bar{y}_{.j2} - \bar{y}_{..2}) & \sum_j n_j (\bar{y}_{.j2} - \bar{y}_{..2})^2 \end{bmatrix} \quad (5.25)$$

donde los elementos de la diagonal principal representan las sumas cuadráticas intergrupos para cada variable dependiente, y los elementos que están fuera de la diagonal hacen referencia a la relación entre ambas variables.

En el caso univariado es posible dividir la suma cuadrática intergrupos entre la suma cuadrática residual y, tras promediarlas en función de sus grados de libertad, llegar a una estimación de la razón F . En el análisis multivariado, sin embargo, la estimación resulta algo más compleja. Así, la matriz análoga al cociente $SC_{\text{entre}}/SC_{\text{error}}$ del ANOVA se obtiene multiplicando la matriz B por la inversa de la matriz W , es decir, mediante el producto BW^{-1} . El resultado de esta multiplicación es una matriz $p \times p$.

Dado que la realización de la prueba de significación a partir de esta matriz supone tener en cuenta p^2 elementos, la hipótesis de nulidad suele contrastarse empleando un procedimiento alternativo. Desde este enfoque alternativo, la prueba de significación multivariante puede llevarse a cabo mediante el denominado *análisis discriminante*.

Las *funciones discriminantes*, F_i , se obtienen a partir de las p variables dependientes del MANOVA, realizando una combinación lineal o suma ponderada de dichas variables.

La primera función discriminante, F_1 , constituye una combinación lineal de las variables dependientes caracterizada por el hecho de que el cociente $SC_{\text{entre}}/SC_{\text{error}}$ adquiere su máximo valor en F_1 . Es decir, en el espacio p -dimensional original se define una nueva dimensión, de tal forma que en ella las diferencias entre los grupos son máximas. Siempre que $k \geq 3$ y $p \geq 2$, es posible hallar una segunda función, F_2 , que explica determinada proporción de varianza no explicada por la primera función discriminante. Esta segunda función maximiza la razón $SC_{\text{entre}}/SC_{\text{error}}$ si las puntuaciones en F_2 no están correlacionadas con las puntuaciones en F_1 . Para k grupos y p variables dependientes, la cantidad de funciones discriminantes es igual al valor menor entre $k - 1$ y p .

Cada una de las funciones discriminantes se expresa del modo siguiente:

$$F_i = E_{H_0} \cdot u_i \quad (5.26)$$

donde:

E_{H_0} = Matriz de errores de estimación bajo el modelo de la hipótesis nula.

u_i = Vector de los pesos de la i -ésima función discriminante. Dicho vector se define mediante la siguiente expresión:

$$u_i = \frac{1}{\sqrt{s_i}} v_i \quad (5.27)$$

donde:

v_i = i -ésimo *autovector* (*eigenvector*) de la matriz BW^{-1} de las variables dependientes.

Los procedimientos para obtener dichos *autovectores* pueden consultarse en Green y Carroll (1976) y en Namboodiri (1984).

$\tilde{s}_i$ = Varianza de error asociada al *autovector* v_i , definida mediante la siguiente expresión:

$$\tilde{s}_i = v_i^T \left(\frac{1}{gl_{\text{error}}} W \right) v_i \quad (5.28)$$

Aplicando las ecuaciones anteriores se pueden calcular F_1 y F_2 para cada sujeto y, a partir del análisis multivariado, hallar las matrices W , T y B para estas funciones. En caso de llevar a cabo tales operaciones nos encontraremos con tres matrices diagonales. Dado que los elementos que están fuera de la diagonal de cada matriz toman el valor cero, cabe deducir que no existe correlación de ningún tipo entre F_1 y F_2 . Por otra parte, los dos elementos de la diagonal de la matriz B corresponden a las sumas de cuadrados intergrupos para cada una de las funciones discriminantes y los de la matriz W , a las sumas cuadráticas intragrupo. Por consiguiente, cuando se calcula el producto BW^{-1} se obtienen los dos elementos de la diagonal que expresan el cociente $SC_{\text{entre}}/SC_{\text{error}}$ para cada función. Es evidente que cocientes de mayor valor reflejan mayores diferencias entre las medias poblacionales de los distintos grupos. La cuestión, en el MANOVA, se centra en saber si los cocientes son de una magnitud o suficientemente grande como para rechazar la hipótesis de nulidad.

Cabe señalar que los elementos de la diagonal de la matriz BW^{-1} obtenida a partir de las funciones discriminantes son los *autovalores* (*eigenvalues*) asociados con los *autovectores* (*eigenvectors*) de dicha matriz para las Y variables dependientes originales. En el análisis multivariado de la varianza, un *autovalor* (*eigenvalue*) no es sino la razón entre la SC_{entre} y la SC_{error} para cada una de las funciones discriminantes. Por tanto, *autovalores* (*eigenvalues*) de mayor magnitud reflejan mayores diferencias entre los grupos en una determinada función. A su vez, cabe apuntar que aunque un *autovalor* (*eigenvalue*) puede obtenerse calculando, en primer lugar, la puntuación de cada sujeto en una determinada función discriminante y hallando, posteriormente, el cociente $SC_{\text{entre}}/SC_{\text{error}}$ para dicha función, también puede estimarse directamente a partir de la matriz BW^{-1} de las variables originales (los procedimientos para obtener los *autovalores* o *eigenvalues* pueden consultarse en Green y Carroll, 1976 y en Namboodiri, 1984).

A partir de los *autovalores* (*eigenvalues*) se calculan las *correlaciones canónicas* (ρ_i) o las correlaciones existentes entre cada función discriminante y las variables dependientes. Así, tenemos que:

$$\rho_i = \sqrt{\frac{\lambda_i}{1 + \lambda_i}} \quad (5.29)$$

Tanto los *autovalores*, o *eigenvalues*, como las *correlaciones canónicas* pueden tomarse como referencia para calcular los **índices o estadísticos que permiten contrastar la hipótesis de nulidad multivariante** mediante diferentes criterios. En tales índices, λ_i representa el i -ésimo *autovalor* de la matriz BW^{-1} para las Y variables dependientes.

Los índices multivariantes de uso más frecuente son la *lambda de Wilks*, la *traza de Pillai-Bartlett*, la *raíz mayor de Roy* y la *traza de Hotelling-Lawley*. Tales índices siguen los criterios que se definen a continuación:

• Criterio de la razón de verosimilitud

La prueba de la hipótesis que se ajusta a este criterio se basa en el cálculo de la *lambda de Wilks*, la cual se obtiene mediante la siguiente expresión:

$$\Lambda = \prod_{i=1}^s \frac{1}{1 + \lambda_i} \quad (5.30)$$

Teniendo en cuenta que $\lambda_i = SC_B/SC_W$ para cada una de las funciones discriminantes, se deduce que $[1/(1 + \lambda_i)] = SC_W/SC_T$. Por tanto, la lambda de Wilks es igual al producto entre las varianzas no explicadas por cada una de las funciones discriminantes. Como cabe deducir de lo anterior, la probabilidad de obtener significación estadística aumenta en la medida en que disminuye el valor de Λ .

• Criterio de la raíz más grande

La prueba de la hipótesis que se ajusta a este criterio se basa en el cálculo de la *raíz mayor de Roy*, la cual se obtiene mediante la siguiente fórmula:

$$R = \frac{\lambda_{\max}}{1 + \lambda_{\max}} \quad (5.31)$$

siendo $\lambda_{\max}$ el mayor de los autovalores obtenidos de la matriz BW^{-1} para las Y variables dependientes.

El estadístico R es igual al cociente SC_B/SC_T para la función discriminante asociada al máximo autovalor, de lo que se deduce que este índice sólo toma en consideración una de las funciones discriminantes.

• Criterio de la traza

La prueba de la hipótesis que se ajusta a este criterio se basa en el cálculo de dos estadísticos, a saber, la *traza de Hotelling-Lawley* y la *traza de Pillai-Bartlett*. El primero de ellos se obtiene mediante la siguiente fórmula:

$$T = \sum_{i=1}^s \lambda_i \quad (5.32)$$

la cual equivale a la suma de los cocientes entre la SC_B y la SC_W para cada una de las funciones discriminantes.

La *traza de Pillai-Bartlett*, por su parte, se obtiene mediante la siguiente expresión:

$$V = \sum_{i=1}^s \frac{\lambda_i}{1 + \lambda_i} \quad (5.33)$$

Puesto que $\lambda_i/1 + \lambda_i$ es igual al cociente SC_B/SC_T para cada función discriminante, la traza de Pillai-Bartlett corresponde a la suma de las varianzas explicadas por cada una de las funciones discriminantes.

La coincidencia entre los cuatro índices anteriores sólo se produce cuando existe una única variable dependiente, ya que en tal caso todos ellos son equivalentes al estadístico F del análisis univariado de la varianza. Si se dispone de más de una variable dependiente, la elección del estadístico multivariado requiere una seria consideración acerca de su robustez y de su potencia estadística. Pascual, García y Frías (1995) afirman que el primer criterio

para tomar una decisión debe ser el de la magnitud de los autovalores, de forma que si la traza se acumula principalmente en el primer autovalor, la prueba más potente es la raíz mayor de Roy. No obstante, Olson (1974) demuestra que la elección de este estadístico resulta muy inadecuada si la traza se reparte proporcionalmente entre todos los autovalores, lo que sucede con mucha frecuencia en ciencias del comportamiento. En tales circunstancias, la mayor potencia corresponde a la traza de Pillai-Bartlett, aunque apenas existen diferencias entre la potencia de dicho índice, la de la lambda de Wilks y la de la traza de Hotelling-Lawley. Por otra parte, la traza de Pillai-Bartlett es la más robusta tanto frente al incumplimiento del supuesto de homogeneidad de las matrices de covarianza como frente a la violación de la normalidad multivariable, siendo la raíz mayor de Roy el estadístico menos robusto (Olson, 1974, 1976; Riba, 1990).

Dado que la lambda de Wilks ha sido y sigue siendo el estadístico más utilizado por los investigadores inmersos en el ámbito de las ciencias del comportamiento (Bray y Maxwell, 1993), presentamos únicamente la prueba de significación basada en dicho índice. Independientemente del estadístico que se utilice, la prueba de la hipótesis nula requiere comparar el valor observado o empírico del estadístico con su distribución muestral bajo la hipótesis nula. La distribución muestral de Λ es muy compleja, por ello se recurre a aproximaciones de dicha distribución muestral. Rao (1951) demostró que realizando una complicada transformación, el estadístico Λ da lugar a una variable cuya distribución se aproxima a la distribución F . Dicha transformación se lleva a cabo mediante la siguiente fórmula:

$$R = \frac{(1 - \Lambda^{1/q}) \left(mq - \frac{1}{2} p(k-1) + 1 \right)}{\Lambda^{1/q} p^{(k-1)}} \quad (5.34)$$

donde:

Λ = Lambda de Wilks.

p = Número de variables dependientes.

k = Número de variables independientes.

$m = N - 1 - 1/2(p + k)$, siendo N = número total de observaciones.

$$q = \frac{p^2(k-1)^2 - 4^{1/2}}{p^2 + (k-1)^2 - 5}$$

La distribución del estadístico R sigue aproximadamente la distribución F con $p(k-1)$ y $mq - 1/2p(k-1) + 1$ grados de libertad. Siempre que $k = 2$ o $k = 3$ (independientemente del valor adoptado por p) o que $p = 1$ o $p = 2$ (independientemente de k), la distribución de R es exactamente igual a la de F .

Pascual, García y Frías (1995) proporcionan las aproximaciones a valores de la distribución F de todos los índices multivariados y presentan varios ejemplos en los que muestran cómo se lleva a cabo la prueba de la hipótesis de nulidad con cada uno de tales estadísticos.

Al igual que en el caso del ANOVA, si el MANOVA nos lleva a rechazar la hipótesis nula multivariada, es necesario examinar entre qué grupos específicos existen diferencias estadísticamente significativas y, a su vez, hemos de analizar qué variables dependientes son las que contribuyen en mayor medida a la discriminación entre los grupos. El estudio de estas cuestiones requiere llevar a cabo nuevos análisis tanto con las variables criterio

(dependientes) como con las de clasificación (independientes). A continuación se describen de forma genérica las **principales estrategias que pueden emplearse tras el rechazo de la hipótesis nula multivariada**.

En relación con el **análisis de las variables criterio**, Bray y Maxwell (1993) destacan tres procedimientos: el contraste de hipótesis específicas univariadas, el análisis discriminante y el análisis «step-down».

- El *contraste de hipótesis específicas univariadas* consiste en llevar a cabo análisis univariados de la varianza con cada una de las p variables dependientes. Según Cramer y Bock (1966), ésta ha sido una de las primeras estrategias recomendadas por los investigadores para interpretar las diferencias entre los grupos. También se conoce como *prueba LSD (Least Significant Difference test)* o como *técnica de la F o de la t protegida* (Bock, 1975; Cooley y Lohnes, 1971; Finn, 1974; Hummel y Sligo, 1971; Spectator, 1977; Wilkinson, 1975). La principal desventaja de este método radica en que no consigue controlar adecuadamente el error de tipo I en el conjunto de las p pruebas univariadas (Bray y Maxwell, 1982; Strahan, 1982). Dicho problema puede solventarse aplicando la corrección de Bonferroni (Harris, 1975), pero tanto la realización de pruebas F univariadas como el uso del procedimiento de Bonferroni suponen una importante pérdida de información, ya que ignoran cualquier posible relación entre las p variables.
- Una segunda estrategia que puede emplearse tras el rechazo de la hipótesis nula multivariada es el *análisis discriminante*. Mediante este análisis se obtienen las combinaciones lineales de las p variables que discriminan en mayor medida los k grupos maximizando la razón entre la varianza intergrupos e intragrupo. Así, el análisis discriminante permite obtener s (el menor valor entre p y $k - 1$) funciones discriminantes, cada una de las cuales constituye una combinación lineal de las p variables que se caracteriza por maximizar la diferencia entre los grupos, siempre que las puntuaciones en cada una de tales funciones no estén correlacionadas con las puntuaciones de cualquiera de las funciones precedentes. En definitiva, este análisis permite predecir los valores de la variable independiente a partir de las puntuaciones que obtienen los sujetos en la variable dependiente y, por tanto, posibilita clasificar a los sujetos en distintos grupos en función de tales puntuaciones. Aunque no vamos a entrar en el tema, cabe señalar que existen distintos métodos para interpretar las funciones discriminantes. Las principales técnicas para realizar dicha interpretación pueden consultarse en los trabajos de Darlington, Weinberg y Walberg (1973), Huberty (1975, 1984), Meredith (1964), Porebski (1966a, 1966b), Stevens (1972), Tatsuoka (1971) y Timm (1975). El análisis discriminante proporciona mucha más información que el contraste de hipótesis específicas univariadas. De hecho, mientras los análisis univariados sólo permiten extraer conclusiones acerca de los efectos de cada variable por separado, el análisis discriminante examina la dimensionalidad subyacente a las variables, las relaciones de las variables con esas dimensiones y las interrelaciones que se establecen entre las variables.
- Por último, existe un tercer método que también se emplea con relativa frecuencia tras rechazar la hipótesis nula en el MANOVA. Se trata del procedimiento conocido como *análisis step-down* (Roy, 1958). Este método puede considerarse como una especie de análisis de covarianza en el que las variables criterio se introducen en un orden específico a fin de examinar la contribución relativa de cada una de ellas a la diferenciación entre los grupos. Es un procedimiento muy útil cuando las p variables pueden

ser ordenadas a priori en función de criterios de carácter teórico (Bock, 1975; Bock y Haggard, 1968; Finn, 1974; Stevens, 1972, 1973).

En lo que respecta al **análisis de las variables de clasificación**, se pueden realizar contrastes específicos que permiten discernir si un determinado grupo difiere de cualquier otro grupo o de una combinación lineal entre varios grupos (Stevens, 1972). Estos contrastes pueden ser *ortogonales* o independientes entre sí, o pueden ser *no ortogonales*, circunstancia que se produce cuando existe correlación entre ellos o cuando cada uno de los grupos consta de un número de sujetos diferente.

Por otra parte, podemos aplicar contrastes *multivariados* o *univariados*. Si se pretenden llevar a cabo comparaciones entre dos o más grupos tomando en consideración el vector de las p variables dependientes, se dispone de diversos estadísticos multivariantes para examinar las diferencias entre los grupos (Green, 1978; Porebski, 1966a; Stevens, 1972). Entre tales estadísticos, el más utilizado es la T^2 de Hotelling. Si, por el contrario, el investigador está interesado en estudiar las diferencias entre los grupos en cada una de las p variables por separado, puede recurrir a las técnicas de comparación múltiple que se emplean en el caso del ANOVA (Scheffé, Tukey, etc.). Es obvio que cuando se adopta este último enfoque, sólo deben examinarse las diferencias específicas entre los grupos en aquellas variables que han llevado a un rechazo de la hipótesis de nulidad univariada.

Para terminar, cabe señalar que además de estudiarlas por separado, también es posible **examinar las variables criterio y las variables de clasificación de forma conjunta**. Normalmente, dicho análisis se realiza mediante el procedimiento que se conoce como el *método de los intervalos de confianza simultáneos de Roy (SCI)*. Esta técnica puede consultarse en los trabajos de Morrison (1976) y de Stevens (1973).

Los textos de Arnau (1990a), Bock (1975), Finn (1974), Harris (1975), Marascuilo y Levin (1983), Morrison (1967), Tatsuoka (1971) y Timm (1975), entre otros, brindan la posibilidad de estudiar de forma exhaustiva el análisis multivariado de la varianza y de profundizar en múltiples cuestiones que aquí no hemos hecho sino esbozar.

En los siguientes capítulos, se abordan las modalidades del diseño experimental clásico utilizadas con mayor frecuencia en el ámbito de las ciencias del comportamiento. Para ello, se toma como referencia básica el criterio taxonómico relacionado con la *técnica de control asociada a la estructura del diseño*. No obstante, dentro de los principales bloques de diseños que cabe distinguir en función de dicho criterio, hemos incluido los modelos de diseño que se derivan de los criterios referidos a la *estrategia empleada para la comparación entre los tratamientos administrados a los sujetos*, a la *cantidad de variables independientes de las que consta el diseño* y a la *configuración completa o incompleta de las unidades experimentales*. A su vez, el bloque correspondiente a los diseños asociados a las distintas estrategias de *equilibración* (Capítulo 8), ha sido ampliado a fin de incluir en dicho bloque el *diseño con covariables*, el cual, como ya se ha señalado, se halla vinculado a la técnica de ajuste estadístico denominada *análisis de la covarianza*. Además, se ha elaborado un capítulo adicional (Capítulo 10) en el que se abordan los *diseños multivariantes* y dos tipos de *diseños estructuralmente incompletos* no descritos en los epígrafes precedentes, a saber, el *diseño de bloques incompletos* y el *diseño factorial fraccionado*. De esta forma, la exposición se divide en cuatro grandes bloques: *diseños intergrupos aleatorios* (Capítulos 6 y 7), *diseños que reducen la varianza de error* (Capítulo 8), *diseños de medidas repetidas* (Capítulo 9) y *otras modalidades de diseño* (Capítulo 10). Como se ha señalado previamente, el objetivo de tal sistematización no es otro que presentar aquellos diseños que se consideran como los de mayor uso en la investigación experimental actual.

6

DISEÑOS UNIFACTORIALES ALEATORIOS

6.1. DISEÑO DE DOS GRUPOS ALEATORIOS

6.1.1. Características generales del diseño de dos grupos aleatorios

El *diseño de dos grupos aleatorios* constituye la estructura más elemental de diseño experimental. Consta de un único factor manipulado y de una variable de respuesta. La variable independiente se manipula de tal forma que se generan dos grupos experimentales o condiciones de tratamiento.

Cuando el investigador trabaja con este tipo de diseño selecciona, aleatoriamente, una muestra de sujetos de una población de origen y, posteriormente, asigna al azar los sujetos de la muestra a cada una de las dos condiciones de tratamiento. La selección aleatoria de la muestra incrementa la validez externa del diseño y la asignación aleatoria incide directamente sobre su validez interna. De hecho, la principal ventaja del diseño de dos grupos aleatorios, con respecto a los diseños no-experimentales, radica en el control obtenido mediante la asignación aleatoria de los sujetos, ya que dicho proceso garantiza la equivalencia de los grupos antes de la aplicación del tratamiento. En consecuencia, cualquier diferencia hallada entre las puntuaciones de ambos grupos en la variable dependiente, tras la intervención experimental, puede atribuirse de forma inequívoca a la acción de la variable manipulada.

Esta modalidad de diseño se caracteriza, bien por la ausencia de tratamiento en uno de los dos grupos, en cuyo caso se denomina *diseño con grupo de control*, o bien por la selección de dos valores diferentes de la variable independiente, estructura que se conoce como *diseño de dos grupos de tratamiento*. Por otra parte, el diseño con grupo de control admite muchas variantes atendiendo, fundamentalmente, a la forma en la que se configura el grupo de control. Así, entre tales modalidades cabe citar el *diseño con grupo placebo*, el *diseño*

con grupo de lista de espera, el diseño con grupo de control acoplado, el diseño con grupo de control sin contacto, etc. Aunque cada uno de estos diseños posee sus propias características asociadas a áreas de investigación específicas, se trata de modelos que apenas difieren en sus aspectos formales.

Desde otra perspectiva, los diseños de dos grupos aleatorios también pueden subdividirse en dos categorías diferentes. Nos estamos refiriendo al criterio taxonómico asociado al procedimiento empleado para seleccionar los niveles del factor manipulado. Así, si los niveles seleccionados agotan toda la población de tratamientos administrables, la variable independiente y el diseño se denominan *de efectos fijos*. En este caso, las conclusiones se limitan a los niveles de la variable independiente seleccionados para el estudio, resultando imposible llevar a cabo cualquier tipo de generalización más allá de dichos niveles. En el segundo caso, la población de tratamientos supera en número a los niveles seleccionados, utilizándose un procedimiento aleatorio para seleccionar tales niveles. En consecuencia, las conclusiones se pueden generalizar a niveles que no se utilizan en el estudio. En estas circunstancias, la variable independiente y el diseño se categorizan como *de efectos aleatorios*. Cabe señalar que la mayoría de las variables independientes que se manejan en el ámbito de las ciencias sociales son de efectos fijos.

Aunque la potencia del diseño de dos grupos aleatorios es escasa, Arnau (1986) considera que constituye un modelo muy adecuado para llevar a cabo investigaciones exploratorias con el objetivo de detectar la posible relación existente entre dos variables. Por tal razón, resulta especialmente útil cuando el objetivo del investigador es examinar nuevos problemas, en áreas en las que no se han realizado trabajos previos. No obstante, toda la información que se obtiene mediante este diseño hace referencia a una sola variable independiente y es evidente que, en la realidad, las variables no actúan de forma aislada. Además, al tratarse de un diseño unifactorial que como técnica de control sólo utiliza la aleatorización, genera una cantidad considerable de varianza de error.

6.1.2. El análisis de datos en el diseño de dos grupos aleatorios

6.1.2.1. Modelo general de análisis

La **prueba de la hipótesis** en este tipo de diseño requiere comparar los rendimientos medios obtenidos, por dos grupos independientes, en la variable de respuesta. Cuando los datos son de naturaleza paramétrica, dicho análisis puede llevarse a cabo utilizando pruebas de contrastes de medias para muestras independientes, tales como la *t de Student* o el *análisis unifactorial de la varianza*. Si, por el contrario, los datos son no paramétricos, se deben aplicar pruebas estadísticas basadas en rangos o en frecuencias. Entre estas pruebas, cabe destacar la *U de Mann-Whitney* y la *ji-cuadrado*. En el presente texto nos centraremos en las pruebas de naturaleza paramétrica. El lector interesado en los análisis no paramétricos dispone de excelentes obras dedicadas exclusivamente a tales análisis, entre las que cabe citar el texto clásico de Siegel (1956) o el libro publicado más recientemente por Pardo y San Martín (1994).

Partiendo de la perspectiva de la comparación de modelos descritos en términos de efectos, Pascual (1995a) presenta el **modelo general de análisis unifactorial de la varianza para el diseño de dos grupos aleatorios**. Como señala el autor, la cuestión principal consiste en saber si los dos grupos difieren entre sí en las puntuaciones medias obtenidas en la variable dependiente. Recordemos que, ante tal cuestión, se pueden plantear dos hipótesis: la *hipótesis*

nula y la hipótesis alternativa. La hipótesis nula (H_0) afirma que no existen diferencias estadísticamente significativas entre los grupos, es decir, que las medias de las puntuaciones de tales grupos son iguales:

$$H_0: \mu_1 = \mu_2 \quad (6.1)$$

En consecuencia, el modelo de predicción bajo la hipótesis nula es el siguiente:

$$y_{ij} = \mu + \varepsilon_{ij} \quad (6.2)$$

donde y_{ij} representa la observación efectuada en la variable dependiente para el sujeto i perteneciente al grupo que recibe el tratamiento j , μ hace referencia a la media global, la cual es constante para todos los datos del experimento, y ε_{ij} es un término residual que refleja el efecto de error asociado al sujeto i bajo el tratamiento j .

Por el contrario, la hipótesis alternativa (H_1) predice la existencia de diferencias estadísticamente significativas entre las puntuaciones medias de los grupos, a saber:

$$H_1: \mu_1 \neq \mu_2 \quad (6.3)$$

En consecuencia, el modelo matemático que subyace a esta predicción se formula en los siguientes términos:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (6.4)$$

donde el término adicional α_j representa el efecto específico del tratamiento j o la diferencia entre la media del grupo en cuestión y la media global. Cabe señalar que en el modelo de efectos aleatorios este término adopta la forma de una variable aleatoria.

El objetivo de la prueba de la hipótesis consiste en esclarecer cuál de los dos modelos es más preciso o se ajusta mejor a los datos. Para ello se utiliza el criterio estadístico F , es decir, se compara la variabilidad intergrupos con la variabilidad intragrupo (la lógica y el proceso analítico para la obtención de la razón F ya han sido descritos en el Epígrafe 5.1 correspondiente al *análisis univariado de la varianza*).

No obstante, es importante señalar que el uso adecuado de este criterio estadístico depende del cumplimiento de una serie de **supuestos relativos a los componentes del modelo estructural del diseño**. Tales supuestos son (para una descripción más detallada ver Balluerka, 1999):

- *Supuesto de medida de la variable dependiente*. La variable dependiente debe medirse al menos en una escala de intervalo.
- *Supuesto de normalidad*. Las observaciones o puntuaciones de la variable dependiente deben seguir una distribución normal: $Y \simeq N(\mu, \sigma_Y^2)$.
- *Supuesto de homogeneidad de las varianzas (homocedasticidad)*. Las varianzas deben ser homogéneas para todos los grupos: $\sigma_i^2 = \sigma_j^2 = \sigma_Y^2, \forall i \neq j$.
- *Supuesto de independencia*. Los errores deben ser independientes entre sí: $\varepsilon_{ij} \simeq N(0, \sigma_\varepsilon^2)$.
- *Supuesto de selección aleatoria de los niveles de tratamiento*. Este supuesto corresponde únicamente al diseño de efectos aleatorios y, como ya lo indica su denominación, hace referencia a la necesidad de seleccionar los niveles de la variable independiente siguiendo un criterio completamente aleatorio.

El cumplimiento del *supuesto de selección aleatoria de los niveles de tratamiento*, en el diseño de efectos aleatorios, se garantiza escogiendo al azar tales niveles. A su vez, la selección aleatoria de los sujetos permite obtener observaciones mutuamente independientes

y, por tanto, errores no correlacionados, asegurando el cumplimiento del *supuesto de independencia de los errores* (Keppel, 1982). Este supuesto es especialmente relevante si se pretende aplicar de forma correcta el análisis de la varianza, ya que su violación afecta considerablemente al nivel de significación y a la potencia de la prueba estadística (Pascual y Camarasa, 1991). Autores como Barcikowsky (1981) y Stevens (1992), entre otros, han demostrado que la dependencia entre las observaciones produce un serio incremento en la tasa de error de tipo I. Vallejo (1986a, 1986b) añade que la violación de la independencia también puede llevarnos a cometer un error de tipo II. Un análisis detallado de las consecuencias derivadas del incumplimiento de este supuesto puede encontrarse en los trabajos de Kenny y Judd (1986) y de Scariano y Davenport (1987).

El *supuesto de normalidad* es el más flexible de todos los supuestos y su incumplimiento tiene poca incidencia sobre el test F (p.e. Glass, Peckham y Sanders, 1972). Existen múltiples procedimientos para comprobar si las observaciones siguen la distribución normal y, además, los paquetes estadísticos más comunes incluyen pruebas para evaluar este requisito. Por ejemplo, el SAS y el SPSS ejecutan el *test de Shapiro-Wilk* cuando el tamaño de la muestra es igual o menor a 50 unidades y la *prueba de Kolmogorov-Smirnov* cuando tenemos más de 50 sujetos. Esta última prueba examina si la distribución se ajusta a la curva normal con varianza σ^2 y media μ . Cuando no se conocen los parámetros de la población, es decir, en el caso tan frecuente en el que la muestra procede de una distribución normal con media y varianza desconocidas, se aplica una corrección a la prueba de Kolmogorov-Smirnov denominada *prueba de Lilliefors* (1967). Si nos encontramos ante una infracción drástica del supuesto de normalidad cuando trabajamos, como en el caso que nos ocupa, con diseños unifactoriales intersujetos, podemos recurrir a alguna técnica de análisis no paramétrica alternativa al análisis de la varianza. Por otra parte, las violaciones severas de la normalidad también pueden corregirse transformando las puntuaciones originales a fin de alcanzar la distribución normal (Emerson y Stoto, 1983; Stevens, 1992).

El *supuesto de homogeneidad de las varianzas* es un supuesto importante desde el punto de vista analítico, ya que su incumplimiento puede afectar tanto a la probabilidad de cometer un error de tipo I, como a la potencia de la prueba estadística. El análisis de la varianza es robusto a la violación moderada de la homogeneidad de varianzas (Winer, 1971). Además, algunos autores consideran que, cuando cada condición o grupo de tratamiento está compuesto por el mismo número de sujetos, la distribución F no resulta seriamente distorsionada, aun cuando existan diferencias considerables entre las varianzas. No obstante, cabe señalar que esta afirmación no es compartida por otros autores, quienes cuestionan la robustez del test F ante la heterogeneidad de las varianzas incluso cuando los tamaños muestrales de cada grupo son iguales (Clinch y Keselman, 1982; Rogan y Keselman, 1977; Tomarken y Serlin, 1986; Wilcox, 1987; Wilcox, Charlin y Thompson, 1986). En lo que la mayoría de los autores parecen estar de acuerdo es en que, ante grupos no equilibrados, niveles moderados de heterogeneidad pueden hacer que el error de tipo I o el α real sea sustancialmente diferente del α nominal fijado a priori por el investigador (p.e. Maxwell y Delaney, 1990; Riba, 1990). Ante la posibilidad de obtener criterios estadísticos sesgados, se han desarrollado diferentes pruebas para la verificación del *supuesto de homocedasticidad*. Entre dichas pruebas cabe destacar la *prueba de Hartley*, la de *Cochran*, la de *Bartlett* y la de *Levene*, siendo esta última la más robusta frente a la violación de la normalidad. Cuando los datos no cumplen el *supuesto de homogeneidad de las varianzas* cabe recurrir a varios procedimientos para corregir la heterocedasticidad, tales como, por ejemplo, el *procedimiento de O'Brien*, la *prueba F conservadora*, la *F* de Brown y Forsythe* (1974) y la *prueba W de Welch* (1951).

6.1.2.2. Ejemplo práctico

Supongamos que un investigador desea examinar el efecto de los cursos de preparación para la maternidad sobre el nivel de ansiedad que genera el parto en las mujeres embarazadas. Predice que las mujeres que asisten a tales cursos manifiestan un menor nivel de ansiedad que aquellas que no asisten. Para comprobar este extremo, selecciona al azar una muestra de 10 mujeres embarazadas, asignando aleatoriamente 5 de ellas a dichos cursos y no aplicando ningún tipo de tratamiento a las 5 restantes. Una semana antes del parto, mide el nivel de ansiedad de las mujeres mediante un cuestionario diseñado a tal efecto. Las puntuaciones obtenidas en el cuestionario se presentan en la siguiente tabla.

TABLA 6.1 Matriz de datos del experimento

A (Asistencia a los cursos)	
a_1 (Sí)	a_2 (No)
5	10
2	13
3	15
1	7
1	8

$$\Sigma Y_1 = 12$$

$$\Sigma Y_2 = 53$$

$$\Sigma Y_1^2 = 40$$

$$\Sigma Y_2^2 = 607$$

$$\bar{Y}_1 = 2,4$$

$$\bar{Y}_2 = 10,6$$

$$s_1^2 = \frac{\Sigma Y_1^2 - (\Sigma Y_1)^2/n}{n-1} = 2,8 \quad s_2^2 = \frac{\Sigma Y_2^2 - (\Sigma Y_2)^2/n}{n-1} = 11,3$$

Teniendo en cuenta que la variable dependiente se mide en una escala de intervalo (supuesto de medida de la variable dependiente), que los sujetos han sido asignados aleatoriamente a las condiciones de tratamiento (supuesto de independencia) y partiendo del hecho de que los datos proceden de una población con distribución normal (supuesto de normalidad), sólo resta comprobar el supuesto de homocedasticidad para poder utilizar como prueba de hipótesis una prueba paramétrica, a saber, la t de Student o el análisis unifactorial de la varianza. En primer lugar, aplicaremos la prueba de Hartley a las puntuaciones de los sujetos a fin de verificar el supuesto de homogeneidad de las varianzas, aunque cabe señalar que la prueba de Levene es la prueba más robusta frente al incumplimiento del supuesto de normalidad (la aplicación de esta prueba puede consultarse al final de este apartado, en el epígrafe referido al análisis de datos mediante el paquete estadístico SPSS 10.0). En segundo lugar calcularemos la t de Student y, por último, desarrollaremos el ANOVA mediante diferentes procedimientos.

• Prueba de Hartley

$$F_{\text{máxima}} = \frac{S_{\text{mayor}}^2}{S_{\text{menor}}^2} \quad (6.5)$$

Bajo el supuesto de que las poblaciones de tratamiento son aproximadamente normales y de que, la distribución de este estadístico se aproxima a una distribución F con k y $n - 1$ grados de libertad para el denominador y para el numerador, respectivamente.

En nuestro ejemplo:

$$F_{\text{máxima}} = \frac{11,3}{2,8} = 4,04 \quad ; \quad F_{\text{máxima}}(\alpha = 0,05, k = 2, n - 1 = 4) = 9,60$$

Puesto que el valor observado de F , $F_{\text{obs.}} = 4,03$, es inferior al valor crítico que delimita la región de rechazo, $F_{\text{crit.}} = 9,60$, se acepta la hipótesis nula de igualdad de varianzas entre ambas muestras.

• t de Student

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2)}{\sqrt{\frac{sc_1 + sc_2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = \frac{(2,4 - 10,6)}{\sqrt{\frac{11,2 + 45,2}{5 + 5 - 2} \left(\frac{1}{5} + \frac{1}{5} \right)}} = \frac{-8,2}{\sqrt{2,82}} = -4,883 \quad (6.6)$$

$$sc_1 = \Sigma Y_1^2 - \frac{(\Sigma Y_1)^2}{n} = 40 - \frac{(12)^2}{5} = 11,2$$

donde:

$$sc_2 = \Sigma Y_2^2 - \frac{(\Sigma Y_2)^2}{n} = 607 - \frac{(53)^2}{5} = 45,2$$

Para $n_1 + n_2 - 2 = 8$ grados de libertad, el valor teórico de t a un nivel de significación de 0,05 es 1,86. Dado que el valor observado de t , $t_{\text{obs.}} = -4,883$, es superior al valor teórico con un porcentaje de error del 5 %, $t_{\text{crit.}} = 1,86$, se rechaza la H_0 . En consecuencia, cabe afirmar que existen diferencias estadísticamente significativas en las puntuaciones obtenidas en la escala de ansiedad, entre las mujeres que asisten a los cursos de preparación para la maternidad y las que no reciben dichos cursos.

• Análisis unifactorial de la varianza

En primer lugar, calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y el estadístico F .

Procedimiento 1

$$SCT = \left[\sum_{j=1}^2 \sum_{i=1}^n Y_{ij}^2 \right] - \left[\frac{1}{2 \cdot n} \left[\sum_{j=1}^2 \sum_{i=1}^n Y_{ij} \right]^2 \right] \quad (6.7)$$

$$SCT = (5)^2 + (2)^2 + \dots + (7)^2 + (8)^2 - (65)^2/10 = 647 - 422,5 = 224,5$$

$$SCA = \left[\frac{1}{n} \sum_{j=1}^2 \left[\sum_{i=1}^n Y_{ij} \right]^2 \right] - \left[\frac{1}{2 \cdot n} \left[\sum_{j=1}^2 \sum_{i=1}^n Y_{ij} \right]^2 \right] \quad (6.8)$$

$$SCA = (12)^2/5 + (53)^2/5 - (65)^2/10 = 28,8 + 561,8 - 422,5 = 168,1$$

$$SCR = SCT - SCA = 224,5 - 168,1 = 56,4, \text{ o bien}$$

$$SCR = \left[\sum_{j=1}^2 \sum_{i=1}^n Y_{ij}^2 \right] - \left[\frac{1}{n} \sum_{j=1}^2 \left[\sum_{i=1}^n Y_{ij} \right]^2 \right] = 647 - 590,6 = 56,4 \quad (6.9)$$

Procedimiento 2

A partir de los datos originales se calculan, en primer lugar, las desviaciones de cada uno de esos datos respecto de la media de su grupo, es decir, los errores de estimación bajo el modelo de la hipótesis alternativa, $e_{ij} = y_{ij} - \bar{y}_{.j}$. Posteriormente, se calcula la suma de cuadrados residual.

Y_1	$(Y_{i1} - \bar{Y}_{.1})$	$(Y_{i1} - \bar{Y}_{.1})^2$	Y_2	$(Y_{i2} - \bar{Y}_{.2})$	$(Y_{i2} - \bar{Y}_{.2})^2$
5	2,6	6,76	10	-0,6	0,36
2	-0,4	0,16	13	2,4	5,76
3	0,6	0,36	15	4,4	19,36
1	-1,4	1,96	7	-3,6	12,96
1	-1,4	1,96	8	-2,6	6,76
		$\Sigma (Y_{i1} - \bar{Y}_{.1})^2 = 11,2$			$\Sigma (Y_{i2} - \bar{Y}_{.2})^2 = 45,2$

Por tanto, la suma de cuadrados del error es:

$$SCR = \sum_{j=1}^2 \sum_{i=1}^n (Y_{ij} - \bar{Y}_{.j})^2 = 11,2 + 45,2 = 56,4 \quad (6.10)$$

En segundo lugar, estimamos los parámetros asociados al efecto del tratamiento, α_j , a fin de calcular la suma de cuadrados intergrupos o correspondiente al tratamiento.

	α_j	α_j^2	n_j	$n_j \alpha_j^2$
$\mu_1 - \mu = 2,4 - 6,5$	-4,1	16,81	5	84,05
$\mu_2 - \mu = 10,6 - 6,5$	4,1	16,81	5	84,05
				$\Sigma n_j \alpha_j^2 = 168,1$

Por tanto, la suma de cuadrados intergrupos es:

$$SCA = \sum_{j=1}^2 n_j \alpha_j^2 = 84,05 + 84,05 = 168,1 \quad (6.11)$$

Por último, la suma cuadrática total es:

$$SCT = SCA + SCR = 168,1 + 56,4 = 224,5 \quad (6.12)$$

o bien,

$$SCT = \sum_{j=1}^2 \sum_{i=1}^n (Y_{ij} - \bar{Y}_{..})^2 = (5 - 6,5)^2 + (2 - 6,5)^2 + \dots + (7 - 6,5)^2 + (8 - 6,5)^2 = 224,5 \quad (6.13)$$

Procedimiento 3: Desarrollo mediante vectores

Para calcular las sumas de cuadrados mediante este procedimiento, debemos calcular los vectores correspondientes a los términos que componen la ecuación estructural del ANOVA, a saber, $y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$, es decir, el *vector Y* (vector de puntuaciones directas), el *vector M* (vector de la media general estimada en la muestra), el *vector A* (vector del efecto correspondiente al tratamiento) y el *vector E_{H1}* (vector de los errores de estimación bajo el modelo de la hipótesis alternativa). A partir de tales vectores se calculan las sumas de cuadrados intergrupos, intragrupo y total.

$$Y = \{Y\} = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 1 \\ 1 \\ 10 \\ 13 \\ 15 \\ 7 \\ 8 \end{bmatrix} \quad (6.14) \quad \text{Vector Y}$$

$$M = \{M\} = \begin{bmatrix} 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \end{bmatrix} \quad (6.15) \quad \text{Vector M}$$

$$A = \{\alpha\} = M_a - M = \begin{bmatrix} 2,4 \\ 2,4 \\ 2,4 \\ 2,4 \\ 2,4 \\ 10,6 \\ 10,6 \\ 10,6 \\ 10,6 \\ 10,6 \end{bmatrix} - \begin{bmatrix} 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \end{bmatrix} = \begin{bmatrix} -4,1 \\ -4,1 \\ -4,1 \\ -4,1 \\ -4,1 \\ 4,1 \\ 4,1 \\ 4,1 \\ 4,1 \\ 4,1 \end{bmatrix} \quad (6.16) \quad \text{Vector A}$$

Vector E_{H1}

Partiendo de la fórmula (6.4):

$$\varepsilon_{ij} = y_{ij} - \mu - \alpha_j = y_{i1} - 6,5 - (-4,1) = y_{i1} - 2,4$$

Sustituyendo las puntuaciones y_{i1} por sus correspondientes valores:

$$\begin{aligned} \varepsilon_{11} &= 5 - 2,4 = 2,6 \\ \varepsilon_{21} &= 2 - 2,4 = -0,4 \end{aligned}$$

$$\begin{aligned}
\varepsilon_{31} &= 3 - 2,4 = 0,6 \\
\varepsilon_{41} &= 1 - 2,4 = -1,4 \\
\varepsilon_{51} &= 1 - 2,4 = -1,4 \\
\varepsilon_{i2} &= y_{i2} - \mu - \alpha_2 = y_{i2} - 6,5 - 4,1 = y_{i2} - 10,6 \\
\varepsilon_{12} &= 10 - 10,6 = -0,6 \\
\varepsilon_{22} &= 13 - 10,6 = 2,4 \\
\varepsilon_{32} &= 15 - 10,6 = 4,4 \\
\varepsilon_{42} &= 7 - 10,6 = -3,6 \\
\varepsilon_{52} &= 8 - 10,6 = -2,6
\end{aligned}$$

Por tanto, el vector E_{H_1} adopta los siguientes valores:

$$E_{H_1} = \{\varepsilon_{H_1}\} = \begin{bmatrix} 2,6 \\ -0,4 \\ 0,6 \\ -1,4 \\ -1,4 \\ -0,6 \\ 2,4 \\ 4,4 \\ -3,6 \\ -2,6 \end{bmatrix} \quad (6.17)$$

Para calcular la suma cuadrática total mediante el procedimiento vectorial, hemos de hallar el vector de los errores de estimación correspondientes al modelo de predicción, bajo la hipótesis nula, $y_{ij} = \mu + \varepsilon_{ij}$. Por tanto, debemos partir de la fórmula (6.2) y calcular el error, a saber, $\varepsilon_{ij} = y_{ij} - \mu$. En este primer diseño calcularemos la suma de cuadrados total, tomando como referencia este vector. No obstante, en el resto de los diseños, únicamente calcularemos los vectores correspondientes a los términos del modelo de predicción bajo la hipótesis alternativa, de manera que la suma cuadrática total será derivada de la fórmula $SCT = SC_{\text{intertratamientos}} + SC_{\text{residual}}$.

$$E_{H_0} = \{\varepsilon_{H_0}\} = \begin{bmatrix} 5 \\ 2 \\ 3 \\ 1 \\ 1 \\ 10 \\ 13 \\ 15 \\ 7 \\ 8 \end{bmatrix} - \begin{bmatrix} 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \\ 6,5 \end{bmatrix} = \begin{bmatrix} -1,5 \\ -4,5 \\ -3,5 \\ -5,5 \\ -5,5 \\ 3,5 \\ 6,5 \\ 8,5 \\ 0,5 \\ 1,5 \end{bmatrix} \quad (6.18)$$

Por tanto, las sumas de cuadrados son:

$$SCA = \{\alpha\}^T \{\alpha\} = [-4,1 \ -4,1 \ -4,1 \ -4,1 \ -4,1 \ 4,1 \ 4,1 \ 4,1 \ 4,1 \ 4,1] \begin{bmatrix} -4,1 \\ -4,1 \\ -4,1 \\ -4,1 \\ -4,1 \\ 4,1 \\ 4,1 \\ 4,1 \\ 4,1 \\ 4,1 \end{bmatrix} = 168,1 \quad (6.19)$$

$$SCR = \{\varepsilon_{H_1}\}^T \{\varepsilon_{H_1}\}$$

$$SCR = [2,6 \quad -0,4 \quad 0,6 \quad -1,4 \quad -1,4 \quad -0,6 \quad 2,4 \quad 4,4 \quad -3,6 \quad -2,6] \begin{bmatrix} 2,6 \\ -0,4 \\ 0,6 \\ -1,4 \\ -1,4 \\ -0,6 \\ 2,4 \\ 4,4 \\ -3,6 \\ -2,6 \end{bmatrix} = 56,4 \quad (6.20)$$

$$SCT = \{\varepsilon_{H_0}\}^T \{\varepsilon_{H_0}\}$$

$$SCT = [-1,5 \quad -4,5 \quad -3,5 \quad -5,5 \quad -5,5 \quad 3,5 \quad 6,5 \quad 8,5 \quad 0,5 \quad 1,5] \begin{bmatrix} -1,5 \\ -4,5 \\ -3,5 \\ -5,5 \\ -5,5 \\ 3,5 \\ 6,5 \\ 8,5 \\ 0,5 \\ 1,5 \end{bmatrix} = 224,5 \quad (6.21)$$

o bien: $SCT = SCA + SCR = 168,1 + 56,4 = 224,5$

Tras calcular las sumas de cuadrados mediante diferentes procedimientos, estimaremos las varianzas y el estadístico F .

Como se ha señalado al abordar el modelo general de análisis unifactorial de la varianza para el diseño de dos grupos aleatorios, el objetivo de la prueba de la hipótesis consiste en comparar la cantidad de veces en las que el ajuste del modelo de predicción bajo la hipótesis alternativa es más preciso que el ajuste del modelo de predicción bajo la hipótesis nula. En este sentido, cuanto mayor es el valor de la razón F , mejor se ajustan los datos al modelo de la hipótesis alternativa. Para calcular el estadístico F , debemos obtener el cociente entre la varianza atribuida a la acción del tratamiento y la varianza residual del modelo de predicción bajo la hipótesis alternativa. Para ello, se divide cada suma de cuadrados entre el número de observaciones que intervienen en el cálculo de las desviaciones, es decir, entre sus correspondientes grados de libertad. El concepto de *grado de libertad* hace referencia a la cantidad de datos de un conjunto que pueden modificarse, sin que dicha modificación afecte a los valores alrededor de los cuales está ocurriendo la variación. Por ejemplo, si se reparten 12 libros entre 12 personas, las 11 primeras personas que escojen un libro tienen posibilidad de elección, pero la persona número 12 no tiene otra opción que quedarse con el último libro. Por tanto, con 12 posibles observaciones sólo existen 11 grados de libertad. En nuestro diseño, como la fuente de variación intergrupos proviene de dos grupos experimentales, los grados de libertad correspondientes a esta fuente de variación adoptan el valor 1.

$$\text{Grados de libertad intergrupos} = k - 1 = 2 - 1 = 1 \quad (6.22)$$

Dado que el valor de la media general, μ , está definido, si los dos grupos disponen del mismo número de observaciones, conociendo la media del primer grupo queda definida la media del segundo:

$$\frac{\mu_{a_1} + \mu_{a_2}}{2} = \mu \Rightarrow \mu_{a_2} = 2\mu - \mu_{a_1}$$

En consecuencia, si tenemos más de dos grupos, la media del último grupo experimental queda determinada si se conocen las medias de los restantes grupos y la media general.

Para calcular los grados de libertad correspondientes al término residual del modelo de predicción bajo la hipótesis alternativa se aplica la siguiente fórmula:

$$\text{grados de libertad intragrupo} = N - k = k(n - 1) = 10 - 2 = 8 \quad (6.23)$$

puesto que en cada uno de los k grupos experimentales existen tantos grados de libertad como número de observaciones menos una.

Por último, el valor de los grados de libertad totales se obtiene restando una unidad al número total de observaciones, a saber:

$$\text{grados de libertad totales} = N - 1 = 10 - 1 = 9 \quad (6.24)$$

que son los grados de libertad correspondientes al modelo de predicción que se deriva de la hipótesis nula.

Una vez conocidas las sumas de cuadrados y los grados de libertad, se obtienen las medias cuadráticas de los componentes intergrupos e intragrupo, y se dividen entre sí para estimar el estadístico F .

$$F = \frac{MC_{\text{entre}}}{MC_{\text{error}}} = \frac{\frac{SC_{\text{entre}}}{gl_{\text{entre}}}}{\frac{SC_{\text{error}}}{gl_{\text{error}}}} \quad (6.25)$$

En nuestro caso:

$$F = \frac{\frac{168,1}{1}}{\frac{56,4}{8}} = \frac{168,1}{7,05} = 23,84$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza (Tabla 6.2).

TABLA 6.2 Análisis unifactorial de la varianza para el diseño de dos grupos aleatorios: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Intertratamientos	$SCA = 168,1$	$k - 1 = 1$	$MCA = \frac{SCA}{k - 1} = 168,1$	$F = \frac{MCA}{MCR}$
Error o residual	$SCR = 56,4$	$k(n - 1) = 8$	$MCR = \frac{SCR}{k(n - 1)} = 7,05$	$F = \frac{168,1}{7,05}$
TOTAL	$SCT = 224,5$	$N - 1 = 9$		$F = 23,84$

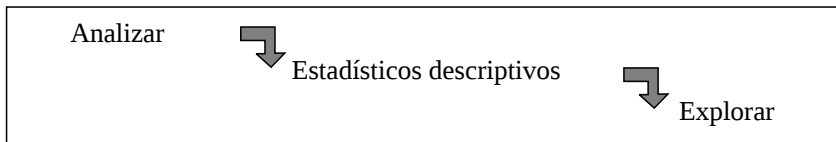
Como la distribución del estadístico F es conocida, se puede determinar cuál es la probabilidad de que, siendo cierta la hipótesis nula, se obtenga una razón de 23,84, teniendo una distribución con 1 y 8 grados de libertad. Dado que la probabilidad de que se obtenga el valor del estadístico $F_{(1,8)} = 23,84$, bajo el modelo de la hipótesis nula, es menor que el valor $\alpha = 0,05$ fijado a priori por el investigador, debemos rechazar el modelo de la hipótesis nula como modelo explicativo de la relación entre los cursos de preparación para la maternidad y el nivel de ansiedad. Para adoptar tal decisión, hemos recurrido a las tablas de los valores críticos de la distribución F , y hemos encontrado que el valor crítico que delimita la región de rechazo, $F_{\text{crít.}(0,05;1,8)} = 5,32$ es menor que el valor observado, $F_{\text{obs.}} = 23,84$, lo que nos lleva a rechazar la hipótesis nula. En consecuencia, cabe afirmar que los cursos de preparación para la maternidad reducen el nivel de ansiedad que genera el parto en las mujeres embarazadas.

• Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

Comprobación de los supuestos del modelo estadístico

- Escogemos la opción **Explorar**



- Indicamos la(s) variable(s) dependiente(s).
- Para el estudio del *supuesto de normalidad*, seleccionamos la opción **Gráficos** y, posteriormente, la opción **Gráficos con pruebas de normalidad**. Esta opción permite calcular la prueba de Kolmogorov-Smirnov con la corrección de Lilliefors para $n > 50$, y la prueba de Shapiro-Wilk para $n \leq 50$.

- Para el estudio del *supuesto de homocedasticidad*, después de introducir la variable independiente en el primer cuadro de diálogo, en la misma opción **Gráficos**, escogemos la opción **Dispersión por nivel con prueba de Levene** y, dentro de este cuadro de diálogo seleccionamos la opción **Estimación de potencia**. Esta opción permite calcular la prueba de Levene. La comprobación del *supuesto de homogeneidad de las varianzas* también puede realizarse desde el menú **Opciones** de los análisis «Comparar medias: Anova de un factor», «MLG: Factorial general» y «MLG: Multivariante».
- La sintaxis para la comprobación del supuesto de normalidad mediante la opción **Explorar** sería:

```
EXAMINE
VARIABLES = ansiedad
/PLOT BOXPLOT NPLOT
/COMPARE GROUP
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

- Resultados:

PRUEBAS DE NORMALIDAD	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Nivel de ansiedad	0,158	10	0,200*	0,926	10	0,431

* Este es un límite inferior de la significación verdadera.

^a Corrección de la significación de Lilliefors.

- La sintaxis para la comprobación del supuesto de homocedasticidad mediante la opción **Explorar** sería:

```
EXAMINE
VARIABLES = ansiedad BY curso
/PLOT BOXPLOT SPREADLEVEL
/COMPARE GROUP
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

- Resultados:

PRUEBA DE HOMOGENEIDAD DE LA VARIANZA		Estadístico de Levene	gl 1	gl 2	Sig.
Nivel de ansiedad	Basándose en la media	3,698	1	8	0,091
	Basándose en la mediana	2,178	1	8	0,178
	Basándose en la mediana y con gl corregido	2,178	1	6,569	0,186
	Basándose en la media recortada	3,579	1	8	0,095

Análisis unifactorial de la varianza

- Escogemos la opción *ANOVA de un factor* del análisis *Comparar medias*.



- Indicamos la(s) variable(s) dependiente(s) y el factor.
- El menú *Opciones* nos ofrece la posibilidad de calcular los estadísticos descriptivos así como de realizar la prueba de homogeneidad de varianzas.
- En nuestro ejemplo, la sintaxis del análisis de la varianza incluyendo la estimación de los estadísticos descriptivos y la comprobación de la homocedasticidad, sería:

ONEWAY
ansiedad BY curso
/STATISTICS DESCRIPTIVES HOMOGENEITY
/MISSING ANALYSIS.

- Resultados:

DESCRIPTIVOS

NIVEL DE ANSIEDAD

	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95 %		Mínimo	Máximo
					Límite inferior	Límite superior		
1,00	5	2,4000	1,6733	0,7483	0,3223	4,4777	1,00	5,00
2,00	5	10,60	3,3615	1,5033	6,4261	14,7739	7,00	15,00
Total	10	6,5000	4,9944	1,5794	2,9272	10,0728	1,00	15,00

PRUEBA DE HOMOGENEIDAD DE VARIANZAS

NIVEL DE ANSIEDAD

Estadístico de Levene	gl 1	gl 2	Sig.
3,698	1	8	0,091

ANOVA

NIVEL DE ANSIEDAD

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	168,100	1	168,100	23,844	0,001
Intra-grupos	56,400	8	7,050		
Total	224,500	9			

6.2. DISEÑO MULTIGRUPOS ALEATORIOS**6.2.1. Características generales del diseño multigrupos aleatorios**

El *diseño multigrupos aleatorios* es una extensión del *diseño de dos grupos aleatorios*. Esta estructura de investigación se caracteriza por el registro de una sola variable dependiente y por la manipulación de un único factor que adopta tres o más niveles de tratamiento. Al igual que en el diseño anterior, los sujetos de la muestra experimental se asignan de forma aleatoria a los distintos niveles de la variable independiente, lo que garantiza la equivalencia entre los diferentes grupos antes de aplicar los tratamientos. En ausencia de estrategias de control estadístico, la equivalencia inicial entre los grupos experimentales constituye un requisito necesario para poder atribuir cualquier diferencia observada, entre sus puntuaciones medias, a la acción de los tratamientos.

La principal ventaja de este diseño, con respecto al diseño de dos grupos aleatorios, radica en que permite obtener una información más precisa acerca de la relación funcional entre la variable independiente y la variable dependiente. Además, cuando la variable manipulada es de naturaleza cuantitativa, resulta posible definir, con gran precisión, el tipo de función matemática que relaciona la variable independiente con la variable de respuesta. Por tal razón, los diseños multigrupos aleatorios se conocen también como *diseños funcionales*. A este respecto, Arnau (1986) señala que, en los casos en los que la relación funcional queda adecuadamente definida, esta modalidad de diseño permite, incluso, interpolar y extrapolar valores que no han sido probados experimentalmente. Así, además de establecer con precisión la relación que existe entre dos variables, los experimentos funcionales posibilitan derivar un modelo cuantitativo exacto que representa la función correspondiente a dicha relación, por lo que resultan muy útiles en el ámbito de la investigación conductual. Sin embargo, poseen las limitaciones propias de los diseños unifactoriales completamente al azar, a saber, generan una cantidad considerable de varianza de error y, por tanto, estiman los efectos con menor precisión que los diseños asociados a técnicas específicas para reducir dicha varianza. Además, como en el caso de los diseños de dos grupos aleatorios, la

manipulación de una sola variable independiente constituye un obstáculo para obtener una representación adecuada de la realidad. Cabe señalar que, al igual que los diseños de dos grupos, los diseños multigrupos aleatorios también pueden categorizarse en función del procedimiento empleado para seleccionar los niveles del factor manipulado. Así, es posible distinguir entre *diseños de efectos fijos* y *diseños de efectos aleatorios*, los cuales son conceptualmente idénticos a los modelos definidos en el apartado anterior (Epígrafe 6.1.1.)

6.2.2. El análisis de datos en el diseño multigrupos aleatorios

6.2.2.1. Posibilidades analíticas para el diseño multigrupos aleatorios

La **prueba de la hipótesis** en este tipo de diseños requiere comparar los resultados obtenidos, en la variable dependiente, por tres o más grupos de sujetos distintos. Cuando los datos son de naturaleza paramétrica y la variable independiente es cualitativa, dicha comparación se lleva a cabo aplicando el *análisis unifactorial de la varianza*. No obstante, cuando los valores de la variable independiente se ordenan de acuerdo con algún criterio cuantitativo y proporcional, resulta más adecuado examinar si existe algún tipo de tendencia que pueda explicar la secuenciación de las distintas medias. Ello es así porque el hecho de conocer qué dirección o qué tendencia presentan los diferentes grupos de datos en función de los valores de la variable independiente aporta mucha más información que la que se obtiene a partir de un simple contraste de medias. En consecuencia, conviene llevar a cabo un *análisis de tendencias*. Una de las técnicas más utilizadas en el ámbito de la psicología, para realizar dicho análisis, consiste en el uso de los *polinomios ortogonales*. Este método permite dividir la variación total debida a los tratamientos, en una serie de componentes que proporcionan información sobre el tipo de tendencia o relación (p.e. lineal, cuadrática, cúbica, etc.) existente entre la variable independiente y la variable dependiente. El objetivo de la prueba consiste en verificar, estadísticamente, cuál es el componente que refleja de manera más adecuada la relación entre estas dos variables.

Por otra parte, si los datos son no paramétricos, se deben aplicar pruebas de significación estadística basadas en la frecuencia o en el orden. Entre tales pruebas, la más utilizada es la que se conoce como *análisis de la varianza unidireccional de Kruskal-Wallis*. No obstante, dicha prueba no permite conocer la posible tendencia que existe entre las distintas condiciones experimentales. Por ello, tras realizar el ANOVA unidireccional de Kruskal-Wallis, se suele aplicar el test denominado *prueba de tendencias de Jonckheere*. Arnau (1986) proporciona varios ejemplos en los que se exponen, de forma muy didáctica, los pasos que deben seguirse tanto para llevar a cabo un análisis de tendencias mediante la técnica de los polinomios ortogonales, como para aplicar el ANOVA unidireccional de Kruskal-Wallis y la prueba de tendencias de Jonckheere.

6.2.2.2. El análisis unifactorial de la varianza

Centrándonos en la solución analítica más común para este tipo de diseños, a saber, en el **análisis unifactorial de la varianza**, cabe señalar que el *modelo estructural* (conocido también como *modelo lineal* o *modelo de efectos*) asociado a la predicción que se realiza bajo la hipótesis alternativa (H_1) es idéntico al del diseño de dos grupos aleatorios. Dicha

ecuación, que representa la puntuación de cualquier sujeto (y_{ij}) como una suma lineal de parámetros poblacionales, adopta la expresión matemática que hemos visto anteriormente:

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (6.4)$$

Como ya se ha indicado en el Apartado 6.1.1, la aplicación adecuada del análisis de la varianza depende del cumplimiento de una serie de supuestos relativos a los componentes de este modelo estructural. Tanto dichos supuestos como los procedimientos para su comprobación y, en caso necesario, para su corrección, pueden consultarse en el citado epígrafe.

6.2.2.3. Ejemplo práctico del análisis unifactorial de la varianza

Supongamos que, en el ámbito de la psicología educativa, nos interesa examinar la influencia que ejercen diferentes estrategias de estudio sobre el recuerdo de un texto. Tras realizar un análisis de las distintas estrategias disponibles para el aprendizaje de textos, escogemos las tres siguientes: (a_1) subrayar el texto, (a_2) realizar esquemas acerca del contenido del texto y (a_3) elaborar una lista de las ideas más importantes del texto. Se selecciona al azar una muestra de 15 sujetos y se asignan, aleatoriamente, 5 de ellos a cada una de las condiciones de tratamiento. Tras el período de aprendizaje, todos los sujetos responden a una prueba de recuerdo (variable criterio). En la Tabla 6.3 puede observarse la cantidad de unidades recordadas por los sujetos en cada una de las condiciones experimentales.

TABLA 6.3 Matriz de datos del experimento

A (Estrategia de estudio)		
a_1 (Subrayado)	a_1 (Esquemas)	a_3 (Listado)
15	30	32
10	32	34
4	25	40
8	27	37
12	35	28
$\Sigma y_1 = 49$ $\Sigma y_1^2 = 549$ $\bar{y}_1 = 9,8$ $S_1^2 = \frac{\Sigma y_1^2 - (\Sigma y_1)^2/n}{n - 1} = 17,2$	$\Sigma y_2 = 149$ $\Sigma y_2^2 = 4.503$ $\bar{y}_2 = 29,8$ $S_2^2 = \frac{\Sigma y_2^2 - (\Sigma y_2)^2/n}{n - 1} = 15,7$	$\Sigma y_3 = 171$ $\Sigma y_3^2 = 5.933$ $\bar{y}_3 = 34,2$ $S_3^2 = \frac{\Sigma y_3^2 - (\Sigma y_3)^2/n}{n - 1} = 21,2$

Dado que la variable dependiente se mide en una escala de intervalo (supuesto de medida de la variable dependiente), que los sujetos han sido asignados aleatoriamente a las condiciones de tratamiento (supuesto de independencia) y suponiendo que los datos proceden de una población con distribución normal (supuesto de normalidad), el único supuesto que queda por comprobar, para poder analizar los datos mediante el análisis unifactorial de la

varianza, es el supuesto de homocedasticidad. Por ello aplicaremos, en primer lugar, la prueba de Hartley a las puntuaciones de los sujetos y, posteriormente, desarrollaremos el ANOVA mediante diferentes procedimientos.

• Prueba de Hartley

$$F_{\text{máxima}} = \frac{S^2_{\text{mayor}}}{S^2_{\text{menor}}} \quad (6.26)$$

Suponiendo que las poblaciones de tratamiento son aproximadamente normales y que $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_j^2$, la distribución de este estadístico se aproxima a una distribución F con k y $n - 1$ grados de libertad para el denominador y para el numerador, respectivamente.

En nuestro ejemplo:

$$F_{\text{máxima}} = \frac{21,2}{15,7} = 1,35 \quad ; \quad F_{\text{máxima}} (\alpha = 0,05, k = 3, n - 1 = 4) = 15,5$$

Puesto que el valor observado de F , $F_{\text{obs.}} = 1,35$, es inferior al valor crítico que delimita la región de rechazo, $F_{\text{crít.}} = 15,5$, se acepta la hipótesis nula de igualdad de varianzas entre los diferentes grupos experimentales.

• Análisis unifactorial de la varianza

Al igual que en el caso del diseño de dos grupos aleatorios (Epígrafe 6.1.2.2), calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y el estadístico F .

Procedimiento 1

1. Suma cuadrática total (SCT).

$$SCT = (15)^2 + (10)^2 + \dots + (37)^2 + (28)^2 - (369)^2/15 = 10.985 - 9.077,4 = 1.907,6$$

2. Suma cuadrática intergrupos (SCA).

$$SCA = (49)^2/5 + (149)^2/5 + (171)^2/5 - (369)^2/15 = 10.768,6 - 9.077,4 = 1.691,1$$

3. Suma cuadrática intragrupo (SCR).

$$SCR = (15)^2 + (10)^2 + \dots + (37)^2 + (28)^2 - [(49)^2/5 + (149)^2/5 + (171)^2/5] = 216,4$$

o bien:

$$SCR = SCT - SCA = 216,4$$

Las fórmulas empleadas en este procedimiento, para el cálculo de las sumas de cuadrados, son las que se expresan a continuación:

$$SCA = \left[\frac{1}{n} \sum_{j=1}^k \left[\sum_{i=1}^n Y_{ij} \right]^2 \right] - \left[\frac{1}{kn} \left[\sum_{j=1}^k \sum_{i=1}^n Y_{ij} \right]^2 \right] \quad (6.27)$$

$$SCR = \left[\sum_{j=1}^k \sum_{i=1}^n Y_{ij}^2 \right] - \left[\frac{1}{n} \sum_{j=1}^k \left[\sum_{i=1}^n Y_{ij} \right]^2 \right] \quad (6.28)$$

$$SCT = \left[\sum_{j=1}^k \sum_{i=1}^n Y_{ij}^2 \right] - \left[\frac{1}{kn} \left[\sum_{j=1}^k \sum_{i=1}^n Y_{ij} \right]^2 \right] \quad (6.29)$$

Téngase en cuenta que si se suman las ecuaciones (6.27) y (6.28), se obtiene la ecuación (6.29). Es decir, la suma entre SCA (suma cuadrática intergrupos) y SCR (suma cuadrática intragrupo) es igual a SCT (suma cuadrática total).

Procedimiento 2

Siguiendo con la misma estrategia empleada en el diseño anterior, obtenemos, en primer lugar, los errores de estimación bajo el modelo de la hipótesis alternativa, $e_{ij} = y_{ij} - \bar{y}_{.j}$. A continuación calculamos la suma de cuadrados intragrupo o residual.

Y_1	$(Y_{i1} - \bar{Y}_{.1})$	$(Y_{i1} - \bar{Y}_{.1})^2$	Y_2	$(Y_{i2} - \bar{Y}_{.2})$	$(Y_{i2} - \bar{Y}_{.2})^2$	Y_3	$(Y_{i3} - \bar{Y}_{.3})$	$(Y_{i3} - \bar{Y}_{.3})^2$
15	5,2	27,04	30	0,2	0,04	32	-2,2	4,84
10	0,2	0,04	32	2,2	4,84	34	-0,2	0,04
4	-5,8	33,64	25	-4,8	23,04	40	5,8	33,64
8	-1,8	3,24	27	-2,8	7,84	37	2,8	7,84
12	2,2	4,84	35	5,2	27,04	28	-6,2	38,44
$\Sigma (Y_{i1} - \bar{Y}_{.1})^2 = 68,8$			$\Sigma (Y_{i2} - \bar{Y}_{.2})^2 = 62,8$			$\Sigma (Y_{i3} - \bar{Y}_{.3})^2 = 84,8$		

Por tanto, la suma de cuadrados residual es:

$$\sum_{j=1}^k \sum_{i=1}^n (Y_{ij} - \bar{Y}_{.j})^2 = 68,8 + 62,8 + 84,8 = 216,4 \quad (6.30)$$

En segundo lugar, estimamos los parámetros asociados al efecto del tratamiento, α_j , a fin de calcular la suma de cuadrados intergrupos o correspondiente al tratamiento.

	α_j	α_j^2	n_j	$n_j \alpha_j^2$
$\mu_1 - \mu = 9,8 - 24,6$	-14,8	219,04	5	1.095,2
$\mu_2 - \mu = 29,8 - 24,6$	5,2	27,04	5	135,2
$\mu_3 - \mu = 34,2 - 24,6$	9,6	92,16	5	460,8
$\Sigma n_j \alpha_j^2$				1.691,2

Por tanto, la suma de cuadrados intergrupos es:

$$SCA = \sum_{j=1}^k n_j \alpha_j^2 = 1.691,2 \quad (6.31)$$

Por último, la suma cuadrática total es:

$$SCT = SCA + SCR = 1.691,2 + 216,4 = 1.907,6 \quad (6.32)$$

o bien:

$$SCT = \sum_{j=1}^k \sum_{i=1}^n (Y_{ij} - \bar{Y}_{..})^2 \quad (6.33)$$

$$SCT = (15 - 24,6)^2 + (10 - 24,6)^2 + \dots + (37 - 24,6)^2 + (28 - 24,6)^2 = 1.907,6$$

Procedimiento 3: Desarrollo mediante vectores

El punto de partida de este procedimiento se halla en la *ecuación estructural* del ANOVA, a saber, en la siguiente ecuación matemática:

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij} \quad (6.4)$$

Para realizar el análisis de la varianza, mediante el procedimiento vectorial, debemos calcular los cuatro vectores siguientes:

$Y = Y_{ij}$ = Vector de puntuaciones directas;

$M = \mu$ = Vector de la media general estimada en la muestra;

$A = \alpha_j$ = Vector del efecto principal del tratamiento;

$E = \varepsilon_{ij}$ = Vector de los residuales o errores.

CÁLCULO:

VECTOR Y

Este vector es el *vector columna* de todas las *puntuaciones directas*.

$$Y = \{Y\} = \begin{bmatrix} 15 \\ 10 \\ 4 \\ 8 \\ 12 \\ 30 \\ 32 \\ 25 \\ 27 \\ 35 \\ 32 \\ 34 \\ 40 \\ 37 \\ 28 \end{bmatrix} \quad (6.34)$$

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada condición experimental para estimar, a partir de tales medias, los parámetros asociados al efecto del tratamiento, α_j .

$$\mu = \frac{1}{k \cdot n} \sum_{j=1}^k \sum_{i=1}^n Y_{ij} = \frac{1}{3 \cdot 5} (15 + 10 + 4 + 8 + 12 + 30 + 32 + 25 + 27 + 35 + 32 + 34 + 40 + 37 + 28) = \dots = \frac{369}{15} = 24,6 \quad (6.35)$$

$$\mu_j = \frac{1}{n} \left[\sum_{i=1}^n Y_{ij} \right] \quad \mu_1 = \frac{1}{5} (15 + 10 + 4 + 8 + 12) = 9,8 \quad (6.36)$$

$$\mu_2 = \frac{1}{5} (30 + 32 + 25 + 27 + 35) = 29,8$$

$$\mu_3 = \frac{1}{5} (32 + 34 + 40 + 37 + 28) = 34,2$$

Por tanto:

$$\alpha_1 = \mu_1 - \mu = 9,8 - 24,6 = -14,8$$

$$\alpha_2 = \mu_2 - \mu = 29,8 - 24,6 = 5,2$$

$$\alpha_3 = \mu_3 - \mu = 34,2 - 24,6 = 9,6$$

El vector A adopta los siguientes valores:

$$A = \{\alpha\} = \begin{bmatrix} -14,8 \\ -14,8 \\ -14,8 \\ -14,8 \\ -14,8 \\ 5,2 \\ 5,2 \\ 5,2 \\ 5,2 \\ 5,2 \\ 9,6 \\ 9,6 \\ 9,6 \\ 9,6 \\ 9,6 \end{bmatrix} \quad (6.37)$$

VECTOR E , CÁLCULO DE LOS RESIDUALES O ERRORES

Partiendo de la fórmula (6.4):

$$\varepsilon_{ij} = y_{ij} - \mu - \alpha_j$$

$$\varepsilon_{i1} = Y_{i1} - \mu - \alpha_1 = Y_{i1} - 24,6 - (-14,8) = Y_{i1} - 9,8$$

Sustituyendo las puntuaciones Y_{i1} por sus correspondientes valores:

$$\varepsilon_{11} = 15 - 9,8 = 5,2$$

$$\varepsilon_{21} = 10 - 9,8 = 0,2$$

$$\varepsilon_{31} = 4 - 9,8 = -5,8$$

$$\varepsilon_{41} = 8 - 9,8 = -1,8$$

$$\varepsilon_{51} = 12 - 9,8 = 2,2$$

$$\varepsilon_{i2} = Y_{i2} - \mu - \alpha_2 = Y_{i2} - 24,6 - 5,2 = Y_{i2} - 29,8$$

Sustituyendo las puntuaciones Y_{i2} por sus correspondientes valores:

$$\varepsilon_{12} = 30 - 29,8 = 0,2$$

$$\varepsilon_{22} = 32 - 29,8 = 2,2$$

$$\varepsilon_{32} = 25 - 29,8 = -4,8$$

$$\varepsilon_{42} = 27 - 29,8 = -2,8$$

$$\varepsilon_{52} = 35 - 29,8 = 5,2$$

$$\varepsilon_{i3} = Y_{i3} - \mu - \alpha_3 = Y_{i3} - 24,6 - 9,6 = Y_{i3} - 34,2$$

Sustituyendo las puntuaciones Y_{i3} por sus correspondientes valores:

$$\varepsilon_{13} = 32 - 34,2 = -2,2$$

$$\varepsilon_{23} = 34 - 34,2 = -0,2$$

$$\varepsilon_{33} = 40 - 34,2 = 5,8$$

$$\varepsilon_{43} = 37 - 34,2 = 2,8$$

$$\varepsilon_{53} = 28 - 34,2 = -6,2$$

Por tanto, el vector E toma los siguientes valores:

$$E = \{\varepsilon\} = \begin{bmatrix} 5,2 \\ 0,2 \\ -5,8 \\ -1,8 \\ 2,2 \\ 0,2 \\ 2,2 \\ -4,8 \\ -2,8 \\ 5,2 \\ -2,2 \\ -0,2 \\ 5,8 \\ 2,8 \\ -6,2 \end{bmatrix} \quad (6.38)$$

Llegados a este punto, podemos calcular la SCA y la SCR aplicando las fórmulas (6.31) y (6.30), respectivamente, o bien multiplicando cada vector por su traspuesto (veáanse las fórmulas (6.39) y (6.40), respectivamente).

Suma cuadrática intergrupos (SCA)

$$SCA = n \cdot \sum_{j=1}^k \alpha_j^2 = 5 [(-14,8)^2 + (5,2)^2 + (9,6)^2] = 1.691,2 \quad (6.31)$$

Suma cuadrática intragrupo (SCR)

$$\begin{aligned} SCR &= \sum_{j=1}^k \sum_{i=1}^n (Y_{ij} - \bar{Y}_{.j})^2 = \sum_{j=1}^k \sum_{i=1}^n \varepsilon_{ij}^2 = (5,2)^2 + (0,2)^2 + (-5,8)^2 + (-1,8)^2 + \\ &+ (2,2)^2 + (0,2)^2 + (2,2)^2 + (-4,8)^2 + (-2,8)^2 + (5,2)^2 + (-2,2)^2 + \\ &+ (-0,2)^2 + (5,8)^2 + (2,8)^2 + (-6,2)^2 = 216,4 \end{aligned} \quad (6.30)$$

La suma entre las dos sumas cuadráticas que acabamos de calcular es igual a la *suma cuadrática total (SCT)*.

Partiendo, por un lado, de los vectores (6.34), (6.37) y (6.38) y, por otro, del vector $\{\mu\}$, podemos volver a formular la ecuación estructural del modelo mediante la siguiente expresión:

$$\{Y\} = \{\mu\} + \{\alpha\} + \{\varepsilon\}$$

Como ya se ha indicado anteriormente, las sumas cuadráticas también pueden calcularse multiplicando cada uno de esos vectores por su vector traspuesto.

Suma cuadrática intergrupos (SCA)

$$SCA = \{\alpha\}^T \cdot \{\alpha\} = 1.691,2$$

$$SCA = (-14,8 \quad -14,8 \quad -14,8 \quad -14,8 \quad -14,8 \quad 5,2 \quad 5,2 \quad 5,2 \quad 5,2 \quad 5,2 \quad 9,6 \quad 9,6 \quad 9,6 \quad 9,6) \begin{bmatrix} -14,8 \\ -14,8 \\ -14,8 \\ -14,8 \\ -14,8 \\ 5,2 \\ 5,2 \\ 5,2 \\ 5,2 \\ 5,2 \\ 9,6 \\ 9,6 \\ 9,6 \\ 9,6 \end{bmatrix} = 1.691,2 \quad (6.39)$$

Como se puede observar, el valor obtenido coincide con el que se deriva de la aplicación de la fórmula (6.31). Aunque este procedimiento pueda parecer más complejo, resulta más fructífero cuando se trabaja con grandes cantidades de datos. Asimismo, cabe señalar que los cálculos que realizan los paquetes estadísticos comparten la lógica de este procedimiento.

La *suma cuadrática intragrupo* se calcula de la siguiente forma:

$$SCR = \{\varepsilon\}^T \cdot \{\varepsilon\} = 216,4$$

$$SCR = (5,2 \quad 0,2 \quad -5,8 \quad -1,8 \quad 2,2 \quad 0,2 \quad 2,2 \quad -4,8 \quad -2,8 \quad 5,2 \quad -2,2 \quad -0,2 \quad 5,8 \quad 2,8 \quad -6,2) \begin{bmatrix} 5,2 \\ 0,2 \\ -5,8 \\ -1,8 \\ 2,2 \\ 0,2 \\ 2,2 \\ -4,8 \\ -2,8 \\ 5,2 \\ -2,2 \\ -0,2 \\ 5,8 \\ 2,8 \\ -6,2 \end{bmatrix} = 216,4 \quad (6.40)$$

Resolviendo esta multiplicación obtenemos el valor que se deriva de la aplicación de la fórmula (6.30).

Independientemente del procedimiento empleado para calcular las sumas de cuadrados, los pasos que se deben seguir a partir de este momento son siempre los mismos. Tras calcular los grados de libertad correspondientes a cada suma de cuadrados, se estiman las varianzas y la razón entre ellas (estadístico *F*).

GRADOS DE LIBERTAD:

$$GL(SCA) = k - 1 = 2$$

$$GL(SCR) = k(n - 1) = 12$$

$$GL(SCT) = kn - 1 = N - 1 = 14$$

(*k*, número de niveles de la variable independiente)

(*n*, número de sujetos por condición de tratamiento) ¹

MEDIAS CUADRÁTICAS (varianzas):

$$MCA = SCA / GL(SCA) = \frac{1.691,2}{2} = 845,6$$

$$MCR = SCR / GL(SCR) = \frac{216,4}{12} = 18,03$$

Llegados a este punto, podemos elaborar la Tabla 6.4 del análisis de la varianza:

TABLA 6.4 Análisis unifactorial de la varianza para el diseño multigrupos aleatorios: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	<i>F</i>
Intertratamientos	<i>SCA</i> = 1.691,2	<i>k</i> - 1 = 2	$MCA = \frac{SCA}{k - 1} = 845,6$	$F = \frac{MCA}{MCR}$
Intragrupo o residual	<i>SCR</i> = 216,4	<i>k</i> (<i>n</i> - 1) = 12	$MCR = \frac{SCR}{k(n - 1)} = 18,03$	$F = \frac{845,6}{18,03}$
Total	<i>SCT</i> = 1.907,6	<i>N</i> - 1 = 14		<i>F</i> = 46,89

Supongamos que establecemos un nivel de confianza del 95 % para la decisión estadística. En este caso, a fin de saber si la *F* observada es estadísticamente significativa, debemos recurrir a las tablas de valores críticos de la distribución *F*_(*α* = 0,05), tomando como grados de libertad para el numerador y para el denominador los valores *gl*₁ = 2 y *gl*₂ = 12, respectivamente. Las tablas nos proporcionan el siguiente valor crítico, *F*_{crít.(0,05;2,12)} = 3,88.

¹ En el modelo de diseño que estamos analizando, tenemos el mismo número de sujetos en todos los niveles de tratamiento.

Dado que $F_{\text{obs.}} > F_{\text{crít.}(0,05;2,12)}$ ($46,89 > 3,88$), podemos rechazar la hipótesis nula (H_0). La probabilidad de error que aceptamos es $p \leq 0,05$. En definitiva, cabe concluir que existen diferencias estadísticamente significativas entre los niveles a_1 , a_2 y a_3 del factor manipulado en la investigación.

6.2.2.4. Comparaciones múltiples entre medias

La razón F es un *test ómnibus de comprobación de hipótesis* o una *prueba de significación general*. En consecuencia, el objetivo de dicha prueba consiste en verificar si las medias de los grupos de tratamiento, consideradas conjuntamente, presentan mayores diferencias de las que cabe esperar por azar. Esta información resulta muy útil para inferir si la variable independiente ejerce influencia sobre la conducta, pero no permite conocer la naturaleza de tal efecto. Por ello, cuando la variable independiente es cualitativa y se obtiene una F significativa, resulta conveniente plantear hipótesis que realicen predicciones más específicas sobre los efectos de la variable independiente. Dado que en el diseño multigrupos aleatorios hay más de dos grupos de tratamiento, existe más de una comparación posible, por ello pueden plantearse diversas hipótesis particulares susceptibles de contrastarse mediante los análisis denominados *contrastes* o *comparaciones múltiples*.

Como señala Pascual (1995a), el número de comparaciones que se llevan a cabo debe estar determinado por dos tipos de razones:

- a) *Razón teórica*: la cantidad y el tipo de contrastes que realiza el investigador deben estar en consonancia con la naturaleza y con los objetivos de la investigación. No se deben llevar a cabo contrastes que carecen de significado teórico.
- b) *Razón de tipo estadístico*: aunque matemáticamente exista un número infinito de contrastes posibles, algunos de ellos son *no ortogonales*, es decir, son redundantes, porque la información que aporta uno de ellos está correlacionada con la que aporta el otro. Como regla general, en toda situación experimental en la que existen k grupos de tratamiento, sólo tenemos $k - 1$ contrastes no redundantes. En consecuencia, resulta conveniente plantear *contrastes ortogonales* que permitan obtener la máxima información realizando el menor número de comparaciones.

Las comparaciones múltiples entre las medias de los grupos pueden ser de dos tipos: *comparaciones planificadas* o *a priori* y *comparaciones no planificadas* o *a posteriori*. Las *comparaciones planificadas* son aquellas que se plantean antes de llevar a cabo el análisis de la varianza y obedecen a intereses que tiene el investigador en la fase previa a la realización del experimento. Las *comparaciones no planificadas*, por el contrario, se formulan en función de los resultados obtenidos en el análisis de la varianza y se llevan a cabo para extraer la máxima información posible de los datos del experimento.

A su vez, las *comparaciones planificadas* se subdividen en dos categorías: *contrastes no ortogonales* y *contrastes ortogonales* o *independientes entre sí*. Como se ha señalado previamente, la diferencia entre ambos tipos de contrastes radica en que estos últimos permiten obtener información no redundante acerca de las posibles diferencias entre las medias de los tratamientos. La principal desventaja de las comparaciones *a priori* es que, a medida que aumenta el número de contrastes, también se incrementa la probabilidad de cometer un error de tipo I o de rechazar la H_0 siendo verdadera. Aunque existen diversos métodos (por ejemplo

la *corrección de Bonferroni* o *prueba de Dunn*) que permiten solventar dicho problema, las comparaciones conocidas como *comparaciones no planificadas* o *a posteriori* constituyen estrategias bastante adecuadas para tal fin, ya que persiguen el objetivo de mantener constante la probabilidad de cometer un error de tipo I en las decisiones estadísticas. Entre dichas estrategias cabe destacar el *método de Fisher*, el *método de Duncan*, el *método de Newman-Keuls*, el *método de Scheffé* y el *método de Tukey* (véase la Figura 6.1).

Además del criterio de clasificación que acabamos de presentar, las técnicas de comparaciones múltiples también pueden categorizarse en función de la cantidad de medias que se incluyen en cada contraste. Así, cuando las condiciones experimentales se comparan dos a dos, nos encontramos ante *comparaciones simples* o *comparaciones entre pares de medias*. Las *comparaciones complejas*, por el contrario, son aquellas comparaciones en las que se trata de comprobar si la media de un grupo difiere de la media de otros dos grupos, o si la media de dos grupos es diferente a la de otros tres, o si difiere de la de cualquier otra posible combinación que incluya más de dos condiciones experimentales.

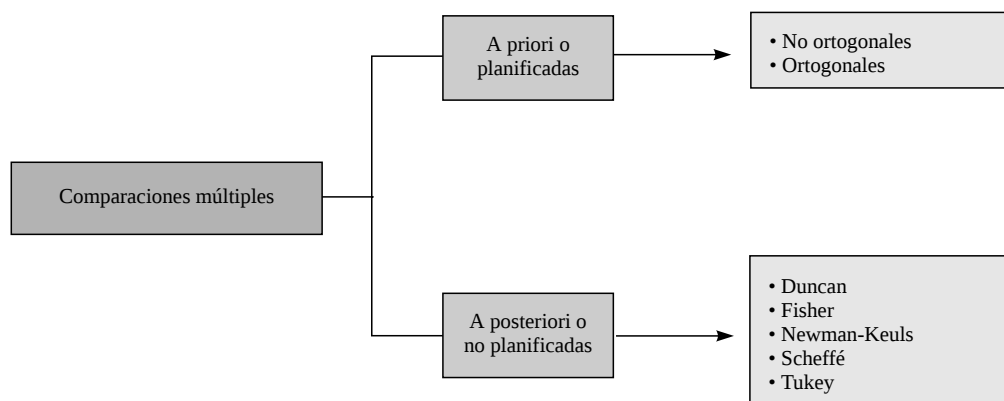


Figura 6.1 Comparaciones múltiples entre medias.

A. Elección del procedimiento para la realización de las comparaciones múltiples

Como se ha señalado anteriormente, a medida que aumenta la cantidad de contrastes que se efectúan, también se incrementa la probabilidad de cometer un error de tipo I, cuestión que debe tenerse en cuenta a la hora de establecer la tasa de error. Una de las dos estrategias básicas de control del α se denomina *tasa de error por comparación* (ERPC *Error Rate Per Comparison*) y consiste en fijar un valor de $\alpha = 0,05$ para cada comparación. Un problema que se deriva de la utilización de esta estrategia es el aumento de la probabilidad de cometer un error de tipo I a medida que aumenta el número de comparaciones. Una segunda estrategia dirigida a controlar el error de tipo I consiste en controlar el valor de α *para un grupo de comparaciones o para todo el experimento*. Dentro de esta estrategia, el investigador puede establecer una *tasa de error por familia* (ERPF *Error Rate Per Family*), según la cual se establece un valor α' (α_{PC}) para cada comparación, de modo que la suma de las tasas de error α' constituye el valor de α . Una segunda manera de controlar el error de tipo I, dentro de la

misma estrategia de control para un grupo de comparaciones o para todo el experimento, consiste en establecer una *tasa de error referida a la familia o al experimento* (ERFW *Error Rate Familywise*), según la cual la probabilidad de cometer un error de tipo I, en un grupo de comparaciones, se define como:

$$\alpha = 1 - (1 - \alpha')^c \quad (6.41)$$

donde:

α = Probabilidad de cometer al menos un error de tipo I al probar c hipótesis nulas o tasa de error referida a la familia o al experimento.

α' = Probabilidad de cometer al menos un error de tipo I al probar una hipótesis nula o tasa de error por comparación.

c = número de comparaciones que se realizan.

Así, en el caso de realizar cuatro comparaciones, si el nivel de $\alpha = 0,05$ y las hipótesis nulas son ciertas, tenemos una probabilidad de cometer, al menos, un error de tipo I de:

$$\alpha = 1 - (1 - 0,05)^4 = 0,1855$$

Como puede deducirse de lo expuesto anteriormente, únicamente la segunda estrategia, a saber, la consistente en controlar el valor de α *para un grupo de comparaciones o para todo el experimento* controla adecuadamente la tasa de error de tipo I. En cuanto a las dos formas de control expuestas dentro de esta estrategia (ERPF y ERFW), aunque difieren en concepto y definición, el control que ejercen es muy similar (Toothaker, 1991). Por este motivo, y para facilitar la exposición, nos referiremos a ambos tipos de control, de forma conjunta, como tasa de error por experimento o α_{PE} . No obstante, debemos señalar que en las tablas que se presentan más adelante (Tablas 6.5 y 6.6), a fin de mantener la máxima exhaustividad y precisión en la descripción de los procedimientos de comparaciones múltiples, se distingue entre los tres tipos de tasas de error.

En términos generales, en la elección del procedimiento más adecuado para llevar a cabo las comparaciones múltiples, deben seguirse dos criterios básicos: 1) que el método escogido controle adecuadamente la tasa de error de tipo I, y 2) que tenga un nivel de potencia adecuado, a saber, que también reduzca la probabilidad de cometer un error de tipo II.

Manteniendo presentes estas premisas y tomando como principal referencia la revisión realizada por Toothaker (1991), expondremos una serie de consideraciones acerca de las principales técnicas de comparaciones múltiples que se utilizan en el ámbito de las ciencias del comportamiento ya que, a nuestro juicio, tales consideraciones pueden resultar muy útiles para decantarse por uno u otro procedimiento (véanse las Tablas 6.5 y 6.6). Posteriormente, desarrollaremos algunos de estos procedimientos tomando como referencia los datos del ejemplo práctico cuyo análisis de la varianza hemos presentado en el subapartado anterior (véase la Tabla 6.3).

En términos generales, siempre y cuando el diseño lo permita, si se aplican *contrastes ortogonales* se controla adecuadamente la tasa de error de tipo I, conservando, a su vez, un alto nivel de potencia.

Diferentes estudios en los que se comparan diversas pruebas de comparaciones múltiples, recogidos en el trabajo de Jaccard, Becker y Wood (1984), demuestran que cuando se trata de técnicas que controlan el $\alpha_{PE} \leq$ nivel nominal, existen pocas diferencias entre los distintos procedimientos en cuanto a su capacidad para maximizar la potencia asociada a una sola comparación (*any-pair power* o *per-pair power*). Estos autores recomiendan la utilización

de la prueba *HSD de Tukey* debido a su simplicidad de cálculo. Cuando se intenta maximizar la potencia para todo el conjunto de comparaciones que se realizan en un experimento (*all-pairs power*), el mejor procedimiento es la *prueba de Peritz*. Asimismo, la prueba de Peritz es la más adecuada cuando se quiere conseguir la máxima potencia en los tres tipos de potencia existentes, a saber, la potencia para cualquier par (*any-pair power*)², la potencia para todos los pares (*all-pairs power*)³ y la potencia para un par (*per-pair power*)⁴. En estos estudios comparativos no se contempla el *método de Bonferroni*, dado que su potencia es menor que la del procedimiento HSD de Tukey cuando se cumplen los supuestos del modelo estadístico (Ury, 1976) y, además, no es una prueba recomendable para realizar todas las posibles comparaciones en diseños intersujetos de efectos fijos. No obstante, al igual que la *prueba de Scheffé*, el método de Bonferroni es muy adecuado en el caso de que el investigador únicamente esté interesado en algunas de las posibles comparaciones o de que quiera llevar a cabo comparaciones complejas. Por otra parte, cabe señalar que si las comparaciones entre pares de medias se establecen entre un grupo de control y el resto de condiciones, el procedimiento más apropiado es el de *Dunnett*.

Por último, es importante destacar que algunos estudios basados en simulaciones *Monte Carlo* parecen confirmar que procedimientos como los de *Newman-Keuls* (Newman, 1939; Keuls, 1952), *Duncan* (1955) o *Fisher* (1935) no controlan adecuadamente la tasa de error nominal (Ramsey, 1981; Seaman, Levin y Serlin, 1991), por lo que su uso no es recomendable.

Aunque nos hemos centrado, fundamentalmente, en los criterios para la elección de las técnicas de comparaciones múltiples de uso habitual en las ciencias del comportamiento, el lector interesado en obtener una perspectiva más amplia de tales técnicas puede consultar el excelente artículo de Jaccard, Becker y Wood (1984) o los textos de Hochberg y Tamhane (1987), Keppel (1982), Klockars y Sax (1986), Winer, Brown y Michels (1991), Toothaker (1991) y Hsu (1996), entre otros.

En las Tablas 6.5 y 6.6 se presentan las técnicas de comparaciones múltiples, utilizadas con mayor frecuencia en el ámbito de las ciencias del comportamiento, para aquellos datos que cumplen y que incumplen, respectivamente, los supuestos del modelo estadístico. En ambas tablas se expone el tipo de distribución que sigue cada una de tales técnicas, el tipo de comparaciones que permite llevar a cabo, su nivel de potencia estadística, el control que ejerce sobre el error de tipo I, así como sus características más relevantes. Cabe destacar que todas las pruebas de comparaciones múltiples que pueden calcularse mediante el paquete estadístico SPSS 10.0 para Windows, se hallan recogidas en estas tablas.

² La «potencia para cualquier par» se define como la probabilidad de detectar cualquier diferencia real entre medias en las *J* medias del experimento (Toothaker, 1991).

³ La «potencia para todos los pares» se define como la probabilidad de detectar todas las diferencias reales entre medias en las *J* medias del experimento (Toothaker, 1991).

⁴ La «potencia para un par» se define como la probabilidad media de detectar una diferencia real entre medias en las *J* medias del experimento (Toothaker, 1991).

TABLA 6.5 Comparaciones múltiples para datos que cumplen los supuestos del modelo estadístico

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
t	Distribución t .	Una sola comparación dos a dos o comparaciones complejas ⁵ .	Superior a todas las pruebas de comparaciones múltiples.	ERPC ⁶ Alta probabilidad de cometer error de tipo I .	
Dunn o Desigualdad de Bonferroni	Distribución t .	Comparaciones planificadas. Comparaciones dos a dos o comparaciones complejas.	Alta para pequeños conjuntos de comparaciones planificadas. Disminuye para grandes conjuntos de comparaciones.	ERPF ⁷ /ERFW ⁸ $\alpha' = \alpha/c$	Se basa en la desigualdad aditiva de Bonferroni.
Dunn-Sidák	Distribución t .	Comparaciones planificadas. Comparaciones dos a dos o comparaciones complejas.	Alta para pequeños conjuntos de comparaciones planificadas ($c < 6$). Disminuye para grandes conjuntos de comparaciones.	ERPF/ERFW $\alpha'' = 1 - (1 - \alpha)^{1/c}$	Se basa en una desigualdad multiplicativa.
HSD de Tukey o prueba de Diferencias Honestamente Significativas	Distribución Q^9 de rango studentizado para J medias.	Comparaciones dos a dos.	Superior a la de las pruebas Dunn y Dunn-Sidák cuando se contrastan todas las posibles comparaciones ($c = J(J - 1)/2$), pero inferior a la de tales pruebas si se realizan menos comparaciones de las posibles y éstas son planificadas. Ello se debe a que la prueba de Tukey se basa en el número de medias y no en el número de comparaciones.	ERFW	

⁵ Las comparaciones complejas son aquellas comparaciones en las que, al menos, uno de los términos de la comparación incluye el promedio entre dos o más condiciones experimentales.

⁶ ERPC (*Error Rate Per Comparison*). Tasa de error por comparación. Se fija un valor de $\alpha = 0.05$ para cada comparación. La probabilidad de cometer un error de tipo I se incrementa a medida que aumenta el número de comparaciones (c).

⁷ ERPF (*Error Rate Per Family*). Tasa de error por familia. Control de α para un grupo de comparaciones. A diferencia de otros tipos de tasas de error que se definen como probabilidades, la tasa de ERPF se define como la media del número de falsos positivos (errores de tipo I) cometidos en un grupo o familia de comparaciones. Se establece un valor de α' , de modo que la suma de estas tasas de error constituye el valor de α ($\alpha \leq c\alpha'$).

⁸ ERFW (*Error Rate Familywise*). Tasa de error referida a la familia. La probabilidad de cometer, al menos, un error de tipo I en un grupo de comparaciones es $\alpha \leq 1 - (1 - \alpha')^c$. Esta forma de control de la tasa de error de tipo I difiere en concepto y definición de la establecida en ERPF, aunque en la práctica el control es muy similar.

⁹ $t = \frac{q}{\sqrt{2}}$.

TABLA 6.5 Comparaciones múltiples para datos que cumplen los supuestos del modelo estadístico (continuación)

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
Newman-Keuls	Distribución Q de rango studentizado.	Comparaciones dos a dos.	Superior potencia que la prueba HSD de Tukey.	El control se sitúa entre ERPC y ERFW: alta probabilidad de cometer error de tipo I .	Método por pasos (stepwise) que sigue una lógica descendente.
REGWQ (Q de Ryan, Einot, Gabriel y Welsch)	Distribución Q de rango studentizado.	Comparaciones dos a dos.	Superior potencia que la prueba HSD de Tukey pero inferior a la de Newman-Keuls.	ERFW	Variante de la prueba de Newman-Keuls.
Scheffé	Distribución F .	Todo tipo de comparaciones (dos a dos, complejas, ortogonales, polinomiales, etc.).	Superior potencia que la prueba de Dunn-Sidák si se realizan más de 8 comparaciones complejas planificadas. No obstante, es una prueba de baja potencia.	ERFW	Prueba muy conservadora que controla adecuadamente el α para todo tipo de comparaciones. Si las comparaciones son complejas y no planificadas, de todas las pruebas expuestas hasta el momento, es la única apropiada para controlar el α utilizando la tasa de error por familia de comparaciones, independientemente de la cantidad de comparaciones que se realicen. Prueba adecuada para experimentos exploratorios (Klockars y Sax, 1986).
F de Newman-Keuls	Distribución F .	Comparaciones dos a dos.	Superior a la de la prueba REGWF.	Control de la tasa de error α para cada grupo de p medias, más que para todas las J medias del experimento. De todos modos, no controla adecuadamente el α .	No es adecuada debido a que se asocia a una alta probabilidad de cometer error de tipo I .
REGWF (F de Ryan, Einot, Gabriel and Welsch)	Distribución F .	Comparaciones dos a dos.	Inferior a la de la prueba F de Newman-Keuls.	ERFW. Mismo α_p que la prueba REGWQ.	Se basa en el mismo estadístico y sigue la misma lógica que la prueba F de Newman-Keuls.

TABLA 6.5 Comparaciones múltiples para datos que cumplen los supuestos del modelo estadístico (continuación)

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
LSD o DMS (Diferencia Menor Significativa) de Fisher	Distribución t .	Comparaciones dos a dos o comparaciones complejas.	Elevada. No se puede establecer la potencia relativa de esta prueba respecto a otras debido a que requiere, en un primer paso, que la prueba F global sea significativa.	No controla la tasa de ERFW en todas las situaciones: Únicamente controla la tasa de ERFW en el primer paso. Posteriormente no se alcanza este tipo de control, ya que todas las comparaciones se llevan a cabo con el mismo valor crítico de t a un mismo nivel α (se toma como referencia el menor valor crítico de t a un nivel α dado para que el resultado sea significativo, considerando una única comparación: Diferencia Menor Significativa).	No requiere tamaños muestrales iguales. Es necesario realizar una prueba F global previa, lo cual hace que esta prueba se conozca como prueba protegida. Sólo se recomienda cuando se prevé que la H_0 es verdadera (Keppel, 1982).
Shaffer-Ryan	Distribución Q de rango studentizado.	Comparaciones dos a dos.	No se puede establecer la potencia relativa de esta prueba respecto a otras debido a que requiere, en un primer paso, que la prueba F global sea significativa.	ERFW	Variante de la prueba REGWQ. Es necesario realizar una prueba F global previa, lo cual hace que esta prueba se conozca como prueba protegida.
Dunnnett	Distribución t .	Comparaciones dos a dos o comparaciones complejas. Puede utilizarse cuando se pretende comparar la media de cada uno de los grupos de tratamiento o el promedio de este conjunto de grupos, con la media de un grupo de control.	Muy elevada cuando se compara la media del grupo de control con cada una de las medias de los restantes grupos.	Controla adecuadamente la tasa de ERFW.	

TABLA 6.5 Comparaciones múltiples para datos que cumplen los supuestos del modelo estadístico (continuación)

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
Peritz	Distribución Q de rango studentizado.	Comparaciones dos a dos.	Superior potencia que la prueba REGWQ, incluso cuando existen múltiples hipótesis nulas.	ERFW	Es una variante del método REGWQ. Su cálculo es complejo y no está incluida en los paquetes estadísticos.
Holm-Shaffer	Distribución t.	Comparaciones dos a dos.	Elevada.	En primer lugar, se contrasta la mayor diferencia entre medias de entre las $(c = J(J - 1)/2)$ posibles comparaciones dos a dos, utilizando el valor crítico de Dunn $\alpha' = \alpha/c$. Si ésta es significativa, la segunda mayor diferencia entre medias se contrasta con $\alpha' = \alpha/(c - 1)$. Para la k -ésima comparación, se utiliza $\alpha' = \alpha/(c - k + 1)$.	Variante de la prueba de Dunn, en la que las comparaciones se ordenan desde la mayor hasta la menor diferencia entre medias. Se puede mejorar esta prueba utilizando una prueba F global en el primer paso. Si la prueba F es significativa, se contrasta la mayor diferencia entre medias con el valor crítico que se utilizaría habitualmente para la segunda mayor diferencia entre medias, y así sucesivamente.
Fisher-Hayter	Distribución Q de rango studentizado.	Comparaciones dos a dos.	Inferior a la de la prueba LSD de Fisher, superior a la de la prueba HSD de Tukey y similar a la de la prueba de Peritz.	Controla el α liberal de la prueba LSD de Fisher. Sugiere utilizar un valor crítico de q , siendo el valor del parámetro de número de medias igual a $J - 1$.	Variante de la prueba LSD de Fisher.

Otras pruebas:

- La prueba denominada Tukeyb o *Wholly Significant Difference* (WSD) es una prueba que utiliza la media de los valores críticos de las pruebas HSD de Tukey y de Newman-Keuls, lo cual implica que sea una prueba menos conservadora que la prueba HSD de Tukey.
- La prueba de Duncan (1955) no se incluye en la tabla, debido a que no controla el α utilizando la tasa de ERFW y a que no ha sido utilizada para desarrollar ninguna otra prueba de comparaciones múltiples adecuada.

TABLA 6.6 Comparaciones múltiples para datos que no cumplen los supuestos del modelo estadístico

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
Tukey-Kramer	Distribución t .	Comparaciones dos a dos.	Elevada en condiciones de homocedasticidad.	Control adecuado de la tasa de ERFW. Se producen valores liberales de α (alta probabilidad de cometer error de tipo I) cuando las varianzas de los grupos son heterogéneas.	Es una variante de la prueba HSD de Tukey. Recomendada para diseños no equilibrados en condiciones de homocedasticidad (Jaccard, Becker y Wood, 1984; Hochberg y Tamhane, 1987). Robusta a la no normalidad. Se puede ver afectada por la presencia de valores extremos (<i>outliers</i>).
Miller-Winer	Distribución t .	Comparaciones múltiples con un grupo de control.	Elevada.	Valores liberales de α (alta probabilidad de cometer error de tipo I). No robusta a la heterocedasticidad.	Segunda variante de la prueba HSD de Tukey. Adecuada para diseños con grupos no equilibrados. No se recomienda su utilización debido a su alta probabilidad de cometer error de tipo I.
Hochberg GT2	Método que relaciona el estadístico t con tamaños muestrales diferentes, con un valor crítico de la distribución de módulo máximo studentizado.	Comparaciones dos a dos.	Menor potencia que la prueba de Tukey-Kramer.	Controla adecuadamente el α cuando las varianzas son homogéneas. No es robusta a la heterocedasticidad cuando los tamaños muestrales son diferentes.	Se basa en la desigualdad de Bonferroni. Adecuada para diseños con grupos no equilibrados.

TABLA 6.6 Comparaciones múltiples para datos que no cumplen los supuestos del modelo estadístico (continuación)

Prueba	Tipo de distribución	Comparaciones	Potencia	Control α	Características
Games-Howell	Distribución de rango studentizado.	Comparaciones dos a dos.	Más potente que el resto de procedimientos, cuando existe heterocedasticidad. Potencia similar a la de la prueba de Tukey-Kramer en caso de existir homocedasticidad (Jaccard, Becker y Wood, 1984).	Única prueba que mantiene el α en un valor próximo a 0,05 utilizando ERFW, para todas las combinaciones entre varianzas (homocedasticidad/heterocedasticidad) y tamaños muestrales (igual/desimal). No controla adecuadamente el error de tipo I cuando los tamaños muestrales son pequeños. Por ello, se recomienda utilizarla para valores $n \geq 6$. Única prueba que controla el α para el total de posibles comparaciones dos a dos, así como para comparaciones dos a dos específicas.	Robusta frente a la heterocedasticidad. Recomendada cuando se incumplen los supuestos del modelo estadístico (Jaccard, Becker y Wood, 1984; Toothaker, 1991). Robusta a la no normalidad. Se puede ver afectada por la presencia de valores extremos (outliers).
C de Dunnett	Distribución de rango studentizado.	Comparaciones dos a dos.	Menor potencia que la prueba de Games-Howell. Mayor potencia que la prueba T3 de Dunnett cuando el n° de G.L. es grande o moderadamente grande ¹⁰ .	Prueba conservadora, cuyo control de α se aproxima al de la prueba de Games-Howell cuando los G.L. se aproximan a infinito. Control más riguroso de α que el de la prueba de Games-Howell en el resto de situaciones.	Robusta frente a la heterocedasticidad. Es una variante de la prueba T2 de Tamhane. Recomendada si se desea mantener un estricto control de α .
T3 de Dunnett	Distribución de módulo máximo studentizado.	Comparaciones dos a dos.	Mayor potencia que la prueba C de Dunnett cuando el n° de G.L. es pequeño ¹⁰ .	Control adecuado de α .	Robusta frente a la heterocedasticidad. Es una variante de la prueba T2 de Tamhane. Recomendada si se desea mantener un estricto control de α .

¹⁰ El punto de corte para decidir si el número de grados de libertad es grande o pequeño depende, entre otros factores, de la razón entre varianzas. Para varianzas iguales o muy similares, un número grande se define como $gl \geq 220$ para $J = 4$ y $gl \geq 440$ para $J = 8$. Para una razón entre varianzas con un valor de 10, grande se define como $gl \geq 52$ para $J = 4$ y $gl \geq 56$ para $J = 8$.

B. Principales procedimientos de comparaciones múltiples: ejemplos prácticos

• Corrección de Bonferroni

Se trata de un procedimiento que permite controlar la tasa de error por experimento cuando las comparaciones múltiples se plantean *a priori*. También se denomina *Desigualdad de Bonferroni* o *prueba de Dunn* (Dunn, 1961; Riba, 1990) y consiste en utilizar como nivel de *alfa* para cada comparación (α_{PC}), el cociente entre el nivel de *alfa* que se quiere asumir en el experimento (α_{PE}) y la cantidad de comparaciones que se realizan (c):

$$\alpha_{PC} = \frac{\alpha_{PE}}{c} \quad (6.42)$$

De este modo, en el caso de formular 5 contrastes, para que la tasa de error por experimento (α_{PE}) se mantenga en 0,05, en cada comparación individual tendremos que asumir el siguiente margen de error:

$$\alpha_{PC} = \frac{\alpha_{PE}}{c} = \frac{0,05}{5} = 0,01$$

y, en este caso, si en cada comparación $\alpha_{PC} = 0,01$, entonces:

$$\alpha_{PE} = 1 - (1 - \alpha_{PC})^c = 1 - (1 - 0,01)^5 = 0,05$$

Como cabe apreciar, aplicando esta corrección, se consigue que la tasa de error por experimento sea igual o inferior al *alfa nominal* de 0,05 que se fija inicialmente.

Retomando los datos del ejemplo práctico utilizado para ilustrar el análisis unifactorial de la varianza en el diseño multigrupos aleatorios (véase la Tabla 6.3), el investigador podría haber diseñado el experimento con el objetivo de efectuar dos contrastes específicos: (1) podría preguntarse si existen diferencias entre los sujetos del grupo que utiliza la estrategia de subrayar el texto y los que realizan esquemas acerca de su contenido y (2) podría preguntarse si la cantidad media de unidades recordadas conjuntamente por el grupo que subraya el texto y por el que realiza esquemas, difiere con respecto a la del grupo que elabora una lista de las ideas más importantes del texto. Veamos cómo se debe proceder en caso de que se planteen tales hipótesis.

Para el primer contraste (hipótesis (1)) planteamos la siguiente hipótesis nula:

$$H_0 \equiv \mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$$

Estimando los valores desconocidos de los parámetros μ , a partir de los estimadores muestrales $\bar{Y}$, bajo la hipótesis nula:

$$\bar{Y}_1 - \bar{Y}_2 = 0$$

$$(1)\bar{Y}_1 + (-1)\bar{Y}_2 + (0)\bar{Y}_3 = 0$$

$$(1 \quad -1 \quad 0) \begin{pmatrix} \bar{Y}_1 \\ \bar{Y}_2 \\ \bar{Y}_3 \end{pmatrix} = 0$$

Siendo C' el vector fila de coeficientes de la matriz de la hipótesis y M_a el vector de medias estimadas en la muestra, la suma de cuadrados explicada por el contraste Ψ será:

$$\Psi = SC_{\Psi} = \frac{(C'M_a)^2}{C'C} \quad (6.43)$$

Aplicando esta fórmula para calcular la suma de cuadrados explicada por el primer contraste, obtenemos:

$$\Psi_1 = SC_{\Psi_1} = \frac{(C'_1 M_a)^2}{C'_1 C_1}$$

$$C'_1 M_a = (1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ -1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0) \begin{bmatrix} 9,8 \\ 9,8 \\ 9,8 \\ 9,8 \\ 9,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 34,2 \\ 34,2 \\ 34,2 \\ 34,2 \\ 34,2 \end{bmatrix} = -100$$

$$C'_1 C_1 = (1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ -1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 10$$

$$\Psi_1 = \frac{(C'_1 M_a)^2}{C'_1 C_1} = \frac{(-100)^2}{10} = \frac{10.000}{10} = 1.000$$

Con el segundo contraste se debe emplear el mismo procedimiento. Recordemos que en este caso se compara el promedio de los dos primeros grupos con el del tercero. Planteemos la hipótesis nula:

$$H_0 \equiv \frac{1}{2} (\mu_1 + \mu_2) = \mu_3 \Rightarrow$$

$$\Rightarrow \frac{1}{2} \mu_1 + \frac{1}{2} \mu_2 - \mu_3 = 0$$

$$\Rightarrow \mu_1 + \mu_2 - 2\mu_3 = 0$$

Expresada mediante estimadores muestrales:

$$\bar{Y}_1 + \bar{Y}_2 - 2\bar{Y}_3 = 0$$

$$(1)\bar{Y}_1 + (1)\bar{Y}_2 + (-2)\bar{Y}_3 = 0$$

La suma de cuadrados correspondiente al segundo contraste es:

$$\Psi_2 = SC_{\Psi_2} = \frac{(C'_2 M_a)^2}{C'_2 C_2}$$

$$C'_2 M_a = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ -2 \ -2 \ -2 \ -2 \ -2) \begin{bmatrix} 9,8 \\ 9,8 \\ 9,8 \\ 9,8 \\ 9,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 29,8 \\ 34,2 \\ 34,2 \\ 34,2 \\ 34,2 \\ 34,2 \end{bmatrix} = -144$$

$$C'_2 C_2 = (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ -2 \ -2 \ -2 \ -2 \ -2) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -2 \\ -2 \\ -2 \\ -2 \\ -2 \end{bmatrix} = 30$$

$$\Psi_2 = \frac{(C'_2 M_a)^2}{C'_2 C_2} = \frac{(-144)^2}{30} = \frac{20.736}{30} = 691,2$$

Una vez calculadas las sumas de cuadrados, podemos realizar el análisis de la varianza de la forma habitual, utilizando como suma de cuadrados del término de error el obtenido en dicho análisis (véase la Tabla 6.4). No obstante, debe tenerse en cuenta que el valor crítico derivado de la corrección de Bonferroni para cada comparación (α_{PC}) es igual a $0,05/2 = 0,025$.

TABLA 6.7 Prueba de la hipótesis del conjunto de dos contrastes entre medias

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F	p
ψ_1	1.000	1	1.000	55,46	< 0,025
ψ_2	691,2	1	691,2	38,34	< 0,025
Error	216,4	12	18,03		

Consultada la tabla de la razón F se comprueba que el valor crítico es $F_{\text{crít.}(0,025; 1,12)} = 6,55$. Por tanto, rechazamos la hipótesis nula tanto para el primer contraste ($F_{\text{obs.}} = 55,46 > F_{\text{crít.}} = 6,55$) como para el segundo ($F_{\text{obs.}} = 38,34 > F_{\text{crít.}} = 6,55$).

Los dos contrastes que hemos planteado en este ejemplo constituyen un sistema de **contrastos ortogonales**, ya que cumplen con los dos requisitos necesarios para que una base de contrastes sea considerada ortogonal:

- 1. El número de contrastes es igual a los grados de libertad intergrupos:

$$c = k - 1 = \text{gl}_{\text{entre}} = 3 - 1 = 2$$

- 2. Cualquier producto entre los coeficientes de los contrastes, dos a dos, es nulo:

$$C_1' C_2 = (1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ -1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 0) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ -2 \\ -2 \\ -2 \\ -2 \\ -2 \end{bmatrix} = 0$$

En el caso de que el investigador no quiera realizar el cálculo de las sumas de cuadrados de los contrastes planificados a priori, puede aplicar el procedimiento de Bonferroni de un modo alternativo. Este método alternativo consiste en determinar el valor de la diferencia mínima entre dos medias que permite rechazar la hipótesis nula manteniendo controlada la tasa de error por experimento. En otras palabras, este método consiste en calcular el *rango crítico entre dos medias*. La diferencia mínima entre dos medias (de los grupos g y h) para poder rechazar la hipótesis nula viene determinada por la siguiente expresión:

$$|\bar{Y}_g - \bar{Y}_h| \geq \sqrt{F_{(\alpha/c, l, gl_{\text{error}})}} \sqrt{MC_{\text{error}} \sum_{j=1}^k \frac{C_j^2}{n_j}} \tag{6.44}$$

Aplicando esta fórmula para contrastar la primera hipótesis nula ($H_0 \equiv \mu_1 = \mu_2$), formulada en los dos contrastes ortogonales, obtenemos:

$$|\bar{Y}_1 - \bar{Y}_2| \geq \sqrt{F_{(0,05/2,1,12)}} \sqrt{18,03 \sum_{j=1}^3 \frac{C_j^2}{n_j}} \Rightarrow \sqrt{6,55} \sqrt{18,03 \left(\frac{1^2}{5} + \frac{-1^2}{5} + \frac{0^2}{5} \right)} = \\ = \sqrt{6,55 \cdot 18,03 \cdot 0,4} = 6,87$$

Según este resultado, la diferencia entre dos medias debe ser, como mínimo, de 6,87 para poder rechazar la hipótesis de igualdad. La diferencia entre el par de medias que estamos comparando es $|\bar{Y}_1 - \bar{Y}_2| = |9,8 - 29,8| = 20$. Como $20 > 6,87$, podemos afirmar que $\mu_1 \neq \mu_2$.

Continuemos con la segunda hipótesis nula planteada en el experimento:

$$\left(H_0 \equiv \mu_3 = \frac{1}{2} (\mu_1 + \mu_2) \right) \\ \left| \bar{Y}_3 - \left(\frac{\bar{Y}_1 + \bar{Y}_2}{2} \right) \right| \geq \sqrt{F_{(0,05/2,1,12)}} \sqrt{18,03 \sum_{j=1}^3 \frac{C_j^2}{n_j}} \Rightarrow \\ \Rightarrow \sqrt{6,55} \sqrt{18,03 \left(\frac{-0,5^2}{5} + \frac{-0,5^2}{5} + \frac{1^2}{5} \right)} = \sqrt{6,55 \cdot 18,03 \cdot 0,3} = 5,95$$

Dado que en este caso la diferencia entre las medias que estamos comparando es:

$$\left| \bar{Y}_3 - \left(\frac{\bar{Y}_1 + \bar{Y}_2}{2} \right) \right| = |34,2 - 19,8| = 14,4 \text{ y } 14,4 > 5,95$$

podemos rechazar la hipótesis de igualdad. En consecuencia, cabe afirmar que:

$$\mu_3 \neq \frac{1}{2} (\mu_1 + \mu_2)$$

• Procedimiento de Dunnett

La prueba desarrollada por Dunnett (1955) es la más potente, si se pretende comparar la media de un determinado grupo, frente a las medias de los restantes grupos. Para aplicar correctamente este procedimiento todas las comparaciones que se lleven a cabo han de ser *simples*, tienen que estar definidas *a priori* y su número debe ser igual a $k - 1$.

La distancia mínima entre dos medias para poder rechazar la hipótesis nula, manteniendo controlada la tasa de error por experimento (α_{PE}), se obtiene mediante la siguiente fórmula:

$$|\bar{Y}_g - \bar{Y}_h| \geq D_{(\alpha, k, g|_{\text{error}})} \sqrt{MC_{\text{error}} \sum_{j=1}^k \frac{C_j^2}{n_j}} \quad (6.45)$$

Supongamos que, en nuestro ejemplo práctico, decidimos comparar la puntuación media del grupo que subraya el texto, tanto con la puntuación media obtenida por los sujetos que realizan esquemas (contraste 1) como con la de los que elaboran una lista de las ideas más importantes del texto (contraste 2). En este ejemplo, el vector C_j para el contraste 1 sería $(1 - 1 0)$, siendo el vector C_j para el segundo contraste $(1 0 - 1)$. El rango crítico entre pares de medias (contraste 1) para poder rechazar la hipótesis nula asumiendo $\alpha_{PE} = 0,05$ es:

$$|\bar{Y}_g - \bar{Y}_h| \geq D_{(0,05,3,12)} \sqrt{18,03 \left(\frac{1^2}{5} + \frac{-1^2}{5} + \frac{0^2}{5} \right)} \Rightarrow 2,50 \sqrt{18,03 \cdot 0,4} = 6,71$$

La diferencia entre los dos primeros grupos debe ser, como mínimo, de 6,71 para rechazar la hipótesis de igualdad entre sus promedios. La diferencia entre el primer par de medias es $|\bar{Y}_1 - \bar{Y}_2| = |9,8 - 29,8| = 20$. Dado que $20 > 6,71$, podemos afirmar que $\mu_1 \neq \mu_2$.

Si utilizamos la fórmula (6.45) para calcular el rango crítico entre la media del primer grupo y la media del tercer grupo (contraste 2), a pesar de que los coeficientes sean $(1 0 - 1)$, debido a que los grupos son equilibrados, dicha distancia es la misma que la calculada para el primer contraste. Como cabe deducir de lo que acabamos de afirmar, cuando se aplica el procedimiento de Dunnett, el cálculo de la distancia entre pares de medias debe realizarse una sola vez, puesto que los coeficientes de los contrastes siempre son iguales a 1 y a -1 , para los grupos que se comparan entre sí, e iguales a 0 para los restantes.

Solamente resulta necesario calcular un rango crítico específico, para cada par de comparaciones, cuando el número de unidades experimentales no es el mismo en todas las condiciones de tratamiento. Dado que en nuestro ejemplo práctico, los tres grupos experimentales constan de la misma cantidad de sujetos, el rango crítico entre pares de medias no sufre variación. Por tanto, como la distancia entre el primer grupo y el tercero es $|\bar{Y}_1 - \bar{Y}_3| = |9,8 - 34,2| = 24,4$ y esta distancia supera el rango crítico (6,71), se rechaza la hipótesis nula de igualdad entre las medias. En consecuencia, cabe concluir que $\mu_1 \neq \mu_3$.

Al igual que en el procedimiento de Bonferroni, en este caso también se pueden calcular las sumas de cuadrados de los contrastes planteados a priori, y obtener un valor crítico con el que puedan compararse los valores de F derivados de cada contraste, siempre y cuando tales contrastes incluyan un solo grupo frente al resto. En caso de aplicar el procedimiento de Dunnett, rechazaremos la hipótesis de igualdad si:

$$F_{\Psi} \geq D_{(z,k,gl_{error})}^2 \quad (6.46)$$

Aplicando esta fórmula a los datos de nuestro ejemplo obtenemos:

$$D_{(z,k,gl_{error})}^2 = D_{(0,05,3,12)}^2 = (2,50)^2 = 6,25$$

que, como cabe observar, es ligeramente inferior al valor crítico obtenido aplicando la corrección de Bonferroni ($F_{(0,05/2,1,12)} = 6,55$).

• Procedimiento DHS de Tukey

El método de las diferencias honestamente significativas de Tukey (1953) es el más potente para realizar todas las comparaciones posibles entre todos los pares de grupos, siempre

y cuando tales comparaciones sean *simples*. En este caso no es necesario definir *a priori* los grupos que se desean comparar entre sí. Cuando tenemos J grupos de tratamiento, el número de comparaciones simples que se pueden llevar a cabo es igual a $J(J - 1)/2$. Dado que todas las comparaciones son dos a dos, en el caso de que el diseño sea equilibrado, debemos calcular una sola vez la distancia crítica entre dos pares de medias. Dicho rango crítico se obtiene mediante la siguiente fórmula:

$$|\bar{Y}_g - \bar{Y}_h| \geq \frac{q_{(\alpha, k, gI_{\text{error}})}}{\sqrt{2}} \sqrt{MC_{\text{error}} \sum_{j=1}^k \frac{C_j^2}{n_j}} \quad (6.47)$$

Aplicando esta fórmula a los datos de nuestro ejemplo:

$$|\bar{Y}_g - \bar{Y}_h| \geq \frac{q_{(0,05,3,12)}}{\sqrt{2}} \sqrt{18,03 \left(\frac{1^2}{5} + \frac{-1^2}{5} + \frac{0^2}{5} \right)} \Rightarrow \frac{3,77}{\sqrt{2}} \sqrt{18,03 \cdot 0,4} = 7,17$$

Para simplificar la interpretación de los datos suele resultar útil elaborar una tabla en la que se calculan las distancias entre todos los pares de medias (véase la Tabla 6.8). Como cabe observar en la Tabla 6.8: $|\bar{Y}_1 - \bar{Y}_2| = 20 > 7,17$, $|\bar{Y}_1 - \bar{Y}_3| = 24,4 > 7,17$ y $|\bar{Y}_2 - \bar{Y}_3| = 4,4 < 7,17$.

TABLA 6.8 Diferencias entre los pares de medias de las tres condiciones experimentales

Grupo		a_1	a_2	a_3
a_1	Subrayado ($\bar{Y}_1 = 9,8$)	0		
a_2	Esquemas ($\bar{Y}_2 = 29,8$)	20	0	
a_3	Listado ($\bar{Y}_3 = 34,2$)	24,4	4,4	0

Por tanto, cabe rechazar la hipótesis nula de igualdad entre los promedios en las dos primeras comparaciones, pero no en la tercera. En consecuencia, podemos concluir que $\mu_1 \neq \mu_2$, $\mu_1 \neq \mu_3$ y $\mu_2 = \mu_3$.

Como en el caso de los dos procedimientos anteriores, podemos calcular los valores empíricos de F , derivados de los contrastes planteados *a priori*, y compararlos con un valor crítico, siempre y cuando las comparaciones incluyan sólo dos grupos. En caso de aplicar el procedimiento de Tukey, rechazaremos la hipótesis de igualdad si:

$$F_{\Psi} \geq \frac{q_{(\alpha, k, gI_{\text{error}})}^2}{2} \quad (6.48)$$

Aplicando esta fórmula a los datos de nuestro ejemplo obtenemos:

$$\frac{q_{(\alpha, k, gI_{\text{error}})}^2}{2} = \frac{q_{(0,05,3,12)}^2}{2} = 3,77$$

que, como cabe observar, es inferior a los valores críticos obtenidos aplicando tanto el procedimiento de Bonferroni ($F_{(0,05/2,1,12)} = 6,55$) como el de Dunnett ($D^2_{(0,05,3,12)} = 6,25$).

• Procedimiento de Scheffé

El procedimiento de Scheffé (1959) es válido para cualquier circunstancia, tanto si se realizan comparaciones *a priori* como *a posteriori* y, tanto si tales comparaciones son *simples* como *complejas*. La distancia crítica entre pares de medias se obtiene mediante la siguiente fórmula:

$$|\bar{Y}_g - \bar{Y}_h| \geq \sqrt{(k-1) F_{(\alpha, k-1, gl_{\text{error}})}} \sqrt{MC_{\text{error}} \sum_{j=1}^k \frac{C_j^2}{n_j}} \quad (6.49)$$

Aplicando esta fórmula a los datos de nuestro ejemplo:

$$|\bar{Y}_g - \bar{Y}_h| \geq \sqrt{2F_{(0,05,2,12)}} \sqrt{18,03 \left(\frac{1^2}{5} + \frac{-1^2}{5} + \frac{0^2}{5} \right)} \Rightarrow \sqrt{2 \cdot 3,88} \sqrt{18,03 \cdot 0,4} = 7,48$$

que, como cabía esperar, nos lleva a las mismas conclusiones que los procedimientos anteriores.

Si en lugar de calcular el rango crítico entre pares de medias, decidimos establecer un valor crítico de F , a fin de compararlo con los valores empíricos de F derivados de los contrastes planteados *a priori*, emplearemos la siguiente fórmula:

$$F_{\Psi} \geq (k-1)F_{(\alpha, k-1, gl_{\text{error}})} \quad (6.50)$$

que, en nuestro caso, adopta el valor de:

$$(3-1)F_{(0,05,2,12)} = 2 \cdot 3,88 = 7,76$$

Como cabe observar, es un valor superior al obtenido con el resto de los procedimientos. Para finalizar, es importante señalar que este método permite controlar el error de tipo I sin restricciones con respecto al número de comparaciones que pueden efectuarse. Sin embargo, es una prueba altamente conservadora.

6.2.2.5. Análisis de datos mediante el paquete estadístico SPSS 10.0

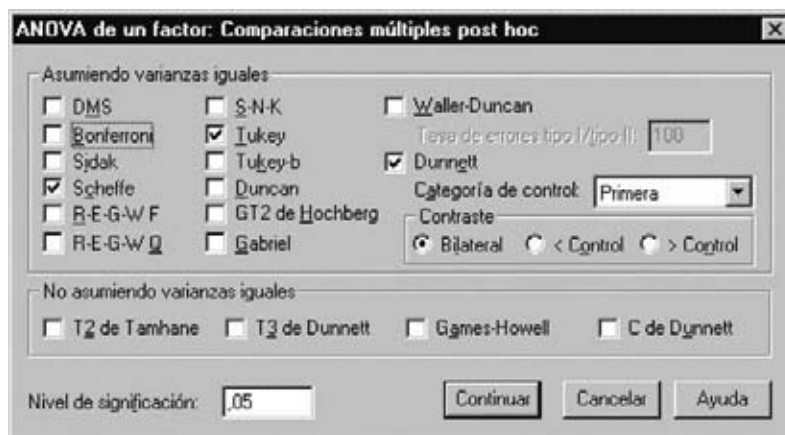
La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

Análisis unifactorial de la varianza

- Escogemos la opción **ANOVA de un factor** del análisis **Comparar medias**.

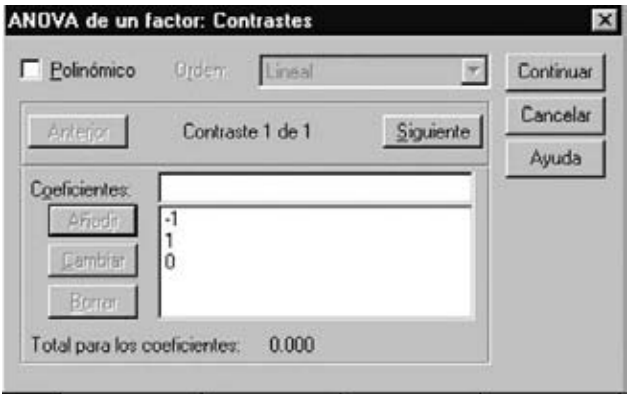


- Indicamos la(s) variable(s) dependiente(s) y el factor.
- El menú **Opciones** nos ofrece la posibilidad de calcular los estadísticos descriptivos así como de realizar la prueba de homogeneidad de varianzas.
- El menú **Post-hoc** proporciona diferentes procedimientos de comparaciones múltiples. Siguiendo con los ejemplos propuestos en el Epígrafe 6.2.2.4, realizaremos las pruebas de Dunnett (seleccionando como categoría de control la primera, es decir, la condición «texto subrayado»), Tukey y Scheffé.



- Para realizar los contrastes planificados mediante la prueba de Dunn o Desigualdad de Bonferroni, escogemos el menú **Contrastes**.

El primer contraste planificado propuesto en el ejemplo práctico del Epígrafe 6.2.2.4 consistía en comparar si existen diferencias entre los sujetos del grupo que utiliza la estrategia de subrayar el texto y los que realizan esquemas acerca de su contenido. Para realizar dicho contraste, indicamos cada uno de los coeficientes de la hipótesis (− 1, 1 y 0, respectivamente) seguido de la opción *Añadir*.



El segundo contraste planificado propuesto en el mismo ejemplo práctico consistía en comparar si la cantidad media de unidades recordadas conjuntamente por el grupo que subraya el texto y por el que realiza esquemas, difiere con respecto a la del grupo que elabora una lista de las ideas más importantes del texto. Para realizar este segundo contraste escogemos, en primer lugar, la opción *Siguiente* y, a continuación, indicamos cada uno de los coeficientes de la hipótesis (− 0,5, − 0,5 y 1, respectivamente) seguido de la opción *Añadir*.

- En nuestro ejemplo, la sintaxis del análisis de la varianza incluyendo la estimación de los estadísticos descriptivos, la comprobación de la homocedasticidad y el cálculo de las pruebas de comparaciones múltiples citadas en el punto anterior, sería:

```
ONEWAY
  recuerdo BY tratam
  /CONTRAST=−1 1 0 /CONTRAST= −0.5 −0.5 1
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /MISSING ANALYSIS
  /POSTHOC = TUKEY SCHEFFE DUNNETT (1) ALPHA(.05).
```

- Resultados:

Descriptivos								
Recuerdo								
	N	Media	Desviación típica	Error típico	Intervalo de confianza para la media al 95 %		Mínimo	Máximo
					Límite inferior	Límite superior		
1	5	9,80	4,15	1,85	4,65	14,95	4	15
2	5	29,80	3,96	1,77	24,88	34,72	25	35
3	5	34,20	4,60	2,06	28,48	39,92	28	40
Total	15	24,60	11,67	3,01	18,14	31,06	4	40

Prueba de homogeneidad de varianzas**Recuerdo**

Estadístico de Levene	gl1	gl2	Sig.
0,049	2	12	0,952

ANOVA**Recuerdo**

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Intergrupos	1.691,200	2	845,600	46,891	0,000
Intragrupos	216,400	12	18,033		
Total	1.907,600	14			

Coefficientes de los contrastes

Contraste	Tratam.		
	1 texto subrayado	2 esquema relacionado	3 ideas principales
1	-1	1	0
2	-0,5	-0,5	1

Pruebas para los contrastes**Recuerdo**

	Contraste	Valor del contraste	Error típico	t	gl	Sig. (bilateral)
Asumiendo igualdad de varianzas	1	20,00	2,69	7,447	12	0,000
	2	14,40	2,33	6,191	12	0,000
No asumiendo igualdad de varianzas	1	20,00	2,57	7,797	7,983	0,000
	2	14,40	2,43	5,936	7,165	0,001

Nota: Como puede constatar el lector, la prueba de contrastes planificados utiliza como prueba de significación estadística la prueba t . Sin embargo, en el ejemplo práctico realizado en el Epígrafe 6.2.2.4., se han calculado los valores de la F de Fisher asociados a cada uno de los contrastes. Cabe recordar que la relación entre ambos estadísticos es: $t = \sqrt{F}$. Retomando los valores obtenidos en el ejemplo práctico, para el primer contraste se obtuvo un valor de $F = 55,46$, correspondiente a un valor de $t = 7,447$. Para el segundo contraste, el valor de F obtenido fue de 38,34, el cual corresponde a un valor de $t = 6,191$.

Pruebas post hoc

Comparaciones múltiples

Variable dependiente: Recuerdo

			Diferencia de medias (I-J)	Error típico	Sig.	Intervalo de confianza al 95 %	
						Límite inferior	Límite superior
HSD de Tukey	(I) tratam.	(J) tratam.					
	Texto subrayado	Esquema relacionado	− 20,00*	2,69	0,000	− 27,17	− 12,83
		Ideas principales	− 24,40*	2,69	0,000	− 31,57	− 17,23
	Esquema relacionado	Texto subrayado	20,00*	2,69	0,000	12,83	27,17
		Ideas principales	− 4,40	2,69	0,268	− 11,57	2,77
	Ideas principales	Texto subrayado	24,40*	2,69	0,000	17,23	31,57
Esquema relacionado		4,40	2,69	0,268	− 2,77	11,57	
Scheffé	Texto subrayado	Esquema relacionado	− 20,00*	2,69	0,000	− 27,49	− 12,51
		Ideas principales	− 24,40*	2,69	0,000	− 31,89	− 16,91
	Esquema relacionado	Texto subrayado	20,00*	2,69	0,000	12,51	27,49
		Ideas principales	− 4,40	2,69	0,298	− 11,89	3,09
	Ideas principales	Texto subrayado	24,40*	2,69	0,000	16,91	31,89
		Esquema relacionado	4,40	2,69	0,298	− 3,09	11,89
t de Dunnett (bilateral) ^a	Esquema relacionado	Texto subrayado	20,00*	2,69	0,000	13,28	26,72
	Ideas principales	Texto subrayado	24,40*	2,69	0,000	17,68	31,12

* La diferencia entre las medias es significativa al nivel 05.
^a Las pruebas t de Dunnett tratan un grupo como control y lo comparan con todos los demás grupos.

Subconjuntos homogéneos

Recuerdo

Tratam.		N	Subconjunto para alfa = 0,05	
			1	2
HSD de Tukey ^a	Texto subrayado	5	9,80	
	Esquema	5		29,80
	relacionado	5		34,20
	Ideas principales		1,000	0,268
Scheffé ^a	Sig.			
	Texto subrayado	5	9,80	
	Esquema	5		29,80
	relacionado	5		34,20
	Ideas principales		1,000	0,298
	Sig.			

Se muestran las medias para los grupos en los subconjuntos homogéneos.
^a Usa tamaño de la muestra de la media armónica = 5,000.



DISEÑOS FACTORIALES ALEATORIOS

7.1. PRINCIPALES CRITERIOS PARA LA CLASIFICACIÓN DE LOS DISEÑOS FACTORIALES

Los *diseños factoriales* son estructuras de investigación en las que se manipulan, simultáneamente, dos o más factores. Tales estructuras pueden categorizarse atendiendo a diferentes criterios, algunos de ellos específicos para este tipo de diseños (criterios de tipo A) y otros comunes al resto de diseños experimentales (criterios de tipo B).

- **Criterios de tipo A**

Entre estos criterios cabe destacar los referidos a la *cantidad de niveles que adoptan las variables independientes* y a la *configuración completa o incompleta de las combinaciones experimentales*.

De acuerdo con el **primer criterio**, las variables que se manipulan en la investigación pueden adoptar la misma cantidad o una cantidad diferente de valores. Cuando la cantidad de niveles de todas las variables independientes es la misma, los diseños se conocen genéricamente como *diseños factoriales de base fija* o *diseños factoriales a k niveles*, y se representan mediante notación exponencial (k^n), en la que el exponente hace referencia a la cantidad de factores de los que consta el diseño y la base, a la cantidad de valores que adopta cada factor. Así, por ejemplo, la potencia 2^3 corresponde a un diseño factorial que consta de tres factores con dos niveles cada uno, es decir, a un diseño factorial a dos niveles. Por otra parte, cuando todas las variables independientes no adoptan la misma cantidad de valores, el diseño se representa mediante un producto (por ejemplo, $p \times q \times r$), donde cada elemento expresa el número de categorías que se seleccionan de cada variable independiente. Por ejemplo, el producto $3 \times 2 \times 4$ corresponde a un diseño que consta de tres variables independientes, con tres, dos y cuatro niveles, respectivamente.

De acuerdo con el criterio referido a la **configuración completa o incompleta de las combinaciones experimentales**, cabe distinguir entre *diseños factoriales completos* y *diseños factoriales incompletos*. En el primer caso, el plan factorial corresponde a un diseño completamente cruzado, en el que los distintos niveles de cada factor se combinan con los distintos niveles del resto de los factores, generando tantas condiciones experimentales como combinaciones posibles entre los factores. En este caso, la cantidad de condiciones experimentales es igual al producto entre los niveles de los factores. Los *diseños factoriales incompletos*, por el contrario, son aquellas estructuras de investigación que carecen de sujetos en algunas combinaciones de tratamientos. En otras palabras, son diseños en los que no se forman todos los grupos que se derivan de la estructura de cruce entre las variables manipuladas (las principales modalidades de diseño que se distinguen en función de este criterio taxonómico ya han sido descritas en el Capítulo 4).

• Criterios de tipo B

Al igual que los diseños unifactoriales o simples, los diseños factoriales también pueden categorizarse tomando como referencia el procedimiento empleado para seleccionar los niveles de los factores manipulados. No obstante, a las dos categorías que se distinguen en función de este criterio en el caso de los diseños simples (*diseños de efectos fijos* y *diseños de efectos aleatorios*), hay que añadir una nueva categoría, a saber, los *diseños de efectos mixtos*. Este tercer modelo de diseño se caracteriza por poseer, al menos, una variable independiente cuyos niveles son de naturaleza arbitraria y una variable independiente cuyos niveles son seleccionados aleatoriamente a partir de la población de niveles que puede adoptar dicha variable.

En lo que respecta a los criterios de clasificación comunes al resto de los diseños experimentales cabe señalar que, los más utilizados en el caso de los diseños factoriales, son los referidos a la *estrategia empleada para la comparación entre los tratamientos administrados a los sujetos* y a la *técnica de control asociada a la estructura del diseño*. No incidiremos en las modalidades de diseño que se distinguen en función de tales criterios porque éstas ya han sido expuestas en el Capítulo 4. Sin embargo, consideramos necesario señalar que, en el presente epígrafe, se aborda el *diseño factorial intergrupos totalmente al azar de efectos fijos*. Como cabe deducir de su denominación, este tipo de diseño requiere que los sujetos sean asignados a los distintos grupos experimentales siguiendo un criterio totalmente aleatorio, lo que garantiza la equivalencia entre tales grupos antes de administrarles los tratamientos.

7.2. VENTAJAS DEL DISEÑO FACTORIAL FRENTE AL DISEÑO UNIFACTORIAL

Los diseños factoriales presentan claras ventajas frente a los diseños unifactoriales o simples. En el presente epígrafe incidiremos, básicamente, en una de tales ventajas, ya que constituye uno de los aspectos distintivos y realmente importantes de esta modalidad de diseño. No obstante, abordaremos brevemente otra serie de ventajas que también se derivan de la utilización del diseño factorial.

La ventaja más importante que presentan los diseños factoriales, frente a los diseños unifactoriales es que, además de permitir el análisis de los *efectos principales*, o de la influencia que ejerce sobre la variable dependiente cada uno de los factores con independencia del resto, también posibilitan examinar los *efectos de interacción* entre tales factores. Los

efectos de interacción hacen referencia a la influencia que ejerce cada variable independiente, teniendo en cuenta los valores que adoptan el resto de variables independientes, es decir, al efecto conjunto de las variables independientes sobre la conducta. Como señala Kazdin (1992), el efecto de interacción es el resultado de la combinación entre los efectos de varios factores, y su presencia indica que la acción de una variable depende del valor (o valores) que adopta(n) otra(s) variable(s). En otras palabras, ante una interacción significativa (por ejemplo, $A \times B$), se debe ser consciente de que una variable independiente (por ejemplo, A) sólo ejerce influencia sobre la variable de respuesta bajo determinadas condiciones (niveles de B) y que, por tanto, su efecto no puede generalizarse a todas las condiciones de tratamiento, lo que disminuye la validez externa del experimento.

Rosenthal y Rosnow (1984) consideran que el estudio de los efectos de interacción es fundamental por tres razones. En primer lugar, porque las interacciones entre varios factores se producen con mucha frecuencia en el ámbito de las ciencias del comportamiento. En segundo lugar, porque muchos de los efectos de interacción que se han detectado en este campo de trabajo poseen gran importancia desde una perspectiva teórica. Y, por último, porque el concepto de interacción no se entiende de forma adecuada. Según Spector (1993), las interacciones reflejan, más adecuadamente que los efectos principales, la verdadera influencia que ejercen los factores sobre la variable dependiente. Por ello, si se encuentra una interacción significativa, los efectos principales no se deben tomar en consideración. En los mismos términos se expresa Pascual (1995a), quien opina que, si en un diseño factorial se rechaza la hipótesis nula de la interacción, los resultados relacionados con los efectos principales pierden importancia sustantiva. En consecuencia, cuando la interacción no se toma como referencia básica para la interpretación de los resultados, se corre el riesgo de incurrir en graves errores de interpretación teórica.

Con respecto a la interpretación de la interacción, Ottenbacher (1991) y Rosnow y Rosenthal (1989) han demostrado que los errores asociados con la interpretación incorrecta de una interacción significativa constituyen una seria amenaza contra la validez de conclusión estadística. Siguiendo a Rosnow y Rosenthal, las interacciones deben interpretarse examinando las puntuaciones residuales o los efectos interactivos de cada celdilla, tras haber sustraído de las medias de cada combinación de tratamientos el efecto constante y los efectos principales (por ejemplo en el caso de un diseño factorial de dos factores: $\alpha\beta_{jk} = \mu_{jk} - \mu - \alpha_j - \beta_k$). Sin embargo, Meyer (1991) no se muestra de acuerdo con esta propuesta, y plantea una interpretación alternativa basada fundamentalmente en los efectos simples, y no en los parámetros usuales de estimación de la interacción. Otros autores, tales como Maxwell y Delaney (1990) y Winer, Brown y Michels (1991), también consideran que la comprobación de las hipótesis asociadas a los efectos simples constituye el procedimiento más adecuado para analizar los efectos de interacción, tras verificar que realmente existen. No obstante, la polémica suscitada en torno a la interpretación de los efectos de interacción llevó a Levin y Marascuilo (1972) a acuñar el concepto de *error de tipo IV*, el cual hace referencia a la interpretación incorrecta que se realiza de los resultados tras rechazar correctamente la hipótesis nula.

Nosotros estamos de acuerdo con Pascual (1995a) en que, cuando la interacción es significativa, los efectos principales no son consistentes y, por tanto, se deben agotar las posibilidades del modelo propuesto. Una manera de lograr tal objetivo consiste en estimar los efectos simples y, si procede, aplicar posteriormente contrastes individuales. Esta es la estrategia de análisis más utilizada entre los metodólogos del comportamiento, quienes, además del problema referido a la interpretación de la interacción, deben afrontar otra cuestión de suma importancia, a saber, la relativa al control de la tasa de error de tipo I que se puede cometer al realizar múltiples contrastes con el objeto de analizar la interacción.

Además de proporcionar la posibilidad de examinar los efectos de interacción entre los factores, los diseños factoriales poseen también otra serie de ventajas importantes. La primera de ellas es que, al introducir varios factores como variables independientes en el diseño, es decir, al *factorizar el diseño*, los efectos asociados a tales factores se sustraen del término de error. En consecuencia, se reduce la varianza de error y se incrementa la potencia de la prueba estadística. Además, el diseño factorial supone un notable ahorro de tiempo y de sujetos con respecto al diseño unifactorial ya que, para conseguir la misma información que se obtiene en un experimento en el que se utiliza un diseño factorial, es necesario llevar a cabo varios experimentos unifactoriales. Por último, cabe señalar que, dada la complejidad de la conducta humana, es lógico suponer que la mayoría de los comportamientos no se hallan determinados por la acción de una sola variable, sino que responden a los efectos de un conjunto de factores. De hecho, en nuestra realidad cotidiana, las variables independientes raramente se presentan aisladas, de ahí que el diseño factorial represente la realidad de manera más adecuada que el diseño unifactorial y que, por tanto, posea mayor validez externa.

7.3. EL ANÁLISIS DE DATOS EN LOS DISEÑOS FACTORIALES ALEATORIOS

7.3.1. Posibilidades analíticas para los diseños factoriales aleatorios

El *análisis factorial de la varianza* es el modelo analítico que se utiliza habitualmente para llevar a cabo la **prueba de la hipótesis** en los diseños factoriales totalmente aleatorios. No obstante, cabe señalar que, si las variables independientes son de naturaleza cuantitativa, los resultados deben analizarse utilizando la *regresión múltiple*. La regresión múltiple es un procedimiento de análisis que permite especificar el tipo de relación funcional lineal que existe entre la variable dependiente y las variables independientes. En el caso de los diseños factoriales, dicha estrategia posibilita conocer qué efectos principales y qué interacciones mantienen una relación estadísticamente significativa con la variable dependiente. Los fundamentos y el desarrollo matemático de este método de análisis pueden consultarse en múltiples textos de metodología, entre los que cabe incluir los de Cohen y Cohen (1975), Draper y Smith (1981), Kerlinger y Pedhazur (1973), Lewis-Beck (1980), Peña Sánchez de Rivera (1987) y Vinod y Ullah (1981).

Centrándonos en el *análisis factorial de la varianza*, cabe señalar que, cuando no existen interacciones significativas, se supone que las variables manipuladas tienen efectos aditivos o independientes. En tal caso, el diseño factorial se ajusta al modelo que se conoce como *modelo aditivo*. Si, por el contrario, las variables independientes actúan conjuntamente, de forma que la predicción de un fenómeno no puede realizarse a partir de la simple adición de los efectos de cada una de tales variables, el diseño factorial se ajusta al modelo denominado *no aditivo* (Keppel, 1982). Dado que, como hemos visto anteriormente, una de las principales cualidades del diseño factorial radica en la posibilidad que brinda para estimar los efectos de interacción, el análisis debe partir de la ecuación estructural asociada a la predicción que se realiza bajo la hipótesis alternativa (H_1) en el modelo no aditivo. Si la prueba de la hipótesis permite rechazar dicho modelo, resulta más adecuado eliminar los términos de interacción y plantear el modelo de efectos aditivos. Con el objetivo de mostrarle al lector el desarrollo de los cálculos necesarios para estimar las razones F que permiten llevar a cabo las pruebas de la hipótesis de nulidad para los parámetros asociados tanto a

los efectos principales como a los efectos de interacción, en la presente exposición se abordan los modelos estructurales no aditivos correspondientes a los diseños factoriales con dos y con tres factores. No obstante, es importante señalar que si el análisis nos llevara a aceptar las hipótesis de nulidad asociadas a los efectos de interacción, el modelo no aditivo no resultaría estadísticamente válido. En tal caso, el modelo aditivo sería mucho más adecuado para explicar los datos empíricos, debido a su carácter parsimonioso y a la exclusión de parámetros cuyos efectos no son estadísticamente significativos. Al hilo de esta afirmación, cabe destacar que la elección inadecuada del modelo puede hacernos incurrir en errores de tipo I y de tipo II, llevándonos a representar de forma incorrecta la realidad¹.

7.3.2. Análisis factorial de la varianza para el diseño factorial $A \times B$

7.3.2.1. Modelo general de análisis

Cuando el diseño consta de dos factores, el modelo matemático subyacente a la predicción que se realiza bajo la hipótesis alternativa adopta la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \quad (7.1)$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable A y el k -ésimo nivel de la variable B .

μ = Media general de la variable dependiente, la cual incluye todos los efectos constantes.

α_j = Efecto principal asociado a la administración del j -ésimo nivel de la variable A , o diferencia entre la media del grupo que recibe dicho tratamiento y la media general.

β_k = Efecto principal asociado a la administración del k -ésimo nivel de la variable B , o diferencia entre la media del grupo que recibe dicho tratamiento y la media general.

$(\alpha\beta)_{jk}$ = Efecto producido por la interacción entre el j -ésimo nivel de A y el k -ésimo nivel de B , con independencia de los efectos de cada tratamiento por separado. Se calcula hallando la diferencia entre la media de los sujetos que reciben dicha combinación de tratamientos y los efectos previos.

ε_{ijk} = Término de error o componente aleatorio del modelo. Expresa la diferencia existente entre la puntuación observada y la puntuación pronosticada basándose en los efectos anteriores, los cuales constituyen la parte sistemática del modelo.

¹ Aunque en el presente texto se ha seguido la perspectiva clásica del contraste de hipótesis, los lectores interesados en profundizar en el enfoque del modelado estadístico y de la comparación de modelos, pueden consultar las excelentes obras de Judd y McClelland (1989) y de Maxwell y Delaney (1990), así como los textos elaborados recientemente por el grupo ModEst (2000a, 2000b). En el segundo de los textos del grupo ModEst se desarrollan modelos de análisis para variables dependientes categóricas, cuestión que, hasta el momento, apenas ha sido abordada en los manuales sobre el diseño experimental.

Como en todo modelo de análisis de la varianza, se asume que este componente tiene una distribución independiente y normal, con media igual a cero y varianza σ^2 , dentro de cada combinación de tratamientos.

Este modelo estructural representa un tipo de combinación lineal y, en consecuencia, parte de la asunción de que cada uno de sus componentes debe ser independiente. Como todo modelo descriptivo, incluye una serie de parámetros que no pueden ser observados directamente y que, por tanto, han de ser estimados a partir de los datos. En la estimación de tales parámetros se parte del supuesto de que los valores de las variables independientes han sido seleccionados arbitrariamente, es decir, se presupone que el diseño es un *modelo de efectos fijos*. Sin embargo, como ya se ha señalado anteriormente, pueden producirse situaciones en las que los valores de las variables independientes se escogen al azar entre una población de valores o tratamientos (*diseños de efectos aleatorios*) o en las que los niveles de uno de los factores se seleccionan arbitrariamente y los niveles de otro se escogen mediante un procedimiento aleatorio (*diseño de efectos mixtos*). Arnau (1986) proporciona las fórmulas que se deben utilizar para derivar los valores esperados de las varianzas (medias cuadráticas) y para calcular las razones *F* en cada uno de tales modelos. En el caso que nos ocupa, a saber, en el modelo de efectos fijos, cuya estructura es idéntica a la del modelo general de la regresión, la mejor estimación de los parámetros se obtiene mediante el *método de los mínimos cuadrados* (véase Ato, López y Serrano, 1981; Balluerka, 1997).

Tomando como *modelo de trabajo* el *modelo completo*, es decir, el modelo que incluye todas las fuentes de variación del diseño y sus interacciones, el análisis de la varianza parte de la descomposición de la variación total en dos componentes: la *variación intergrupos* y la *variación intragrupo*. En el diseño factorial de dos factores, la variación intergrupos se descompone, a su vez, en la variación debida al factor *A*, la debida al factor *B* y la que se deriva de la interacción entre ambos factores. La variación intragrupo, por su parte, incluye únicamente la suma cuadrática del componente residual del modelo. A partir de las sumas de cuadrados y de sus correspondientes grados de libertad se obtienen las medias cuadráticas del factor *A*, del factor *B* y de la interacción $A \times B$. Cada una de tales varianzas se contrasta con el mismo término de error, a saber, con la media cuadrática del error. De este modo, se estiman las razones *F* que permiten llevar a cabo las pruebas de la hipótesis de nulidad para cada uno de los parámetros del modelo. Al igual que en el caso de los diseños unifactoriales, cuando en el diseño factorial se rechaza la hipótesis nula, se deben aplicar *contrastes específicos* para determinar entre qué medias existen diferencias estadísticamente significativas. Tales comparaciones múltiples se llevan a cabo mediante los mismos procedimientos que en los diseños unifactoriales. Sin embargo, es necesario tener presente que cuando se realizan dichos contrastes en el modelo con interacción, la cantidad de grupos que se comparan es igual al número de combinaciones entre los niveles de los factores, a saber, al número de «celdillas».

7.3.2.2. Ejemplo práctico

Supongamos que se realiza una investigación para examinar la influencia que ejercen una serie de variables sociolingüísticas sobre el *nivel de competencia que presenta un sujeto bilingüe en su segunda lengua*. En concreto, nos interesa estudiar los posibles efectos del *nivel de competencia de los padres en dicha lengua* (factor *A*) y del *estadio lingüístico o*

*tipología lingüística del niño*² (factor *B*) sobre la *competencia que éste presenta en su segunda lengua*. A tal fin, asignamos aleatoriamente a tres grupos distintos, 8 niños pertenecientes a cada una de las siguientes categorías: (a_1) *el único progenitor competente en dicha lengua es la madre*, (a_2) *el único progenitor competente en dicha lengua es el padre* y (a_3) *ambos progenitores, madre y padre, son competentes en dicha lengua*. A su vez, dividimos cada uno de los grupos en dos subgrupos, en función de las categorías que adopta el factor *B*, a saber: (b_1) *estadío precoz*, cuando el aprendizaje de la segunda lengua se produce antes de los 3 años y (b_2) *estadío tardío*, cuando dicha lengua se adquiere a partir de los 5 años. Siguiendo los criterios citados, se forman 6 grupos de 4 niños ($n = 4$) con edades comprendidas entre los 9 y los 14 años. Tras la formación de los grupos, se mide el nivel de competencia que muestran los sujetos en la segunda lengua (variable criterio). En la Tabla 7.1 pueden observarse los resultados obtenidos en la investigación.

La Tabla 7.1 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan las *dos categorías del factor B* (b_1 y b_2) y en las columnas los *tres niveles del factor A* (a_1 , a_2 y a_3). Por tanto, $k = 1, 2$ y $j = 1, 2, 3$.

TABLA 7.1 Matriz de datos del experimento

		A (Nivel de competencia de los padres en L2)			Medias marginales
		a_1 (Madre competente)	a_2 (Padre competente)	a_3 (Ambos competentes)	
B (Estadío lingüístico)	b_1 (E. precoz)	7	8	13	
		9	11	15	
		10	10	11	
		8	6	10	
	b_2 (E. tardío)	$\bar{Y}_{11} = 8,5$	$\bar{Y}_{21} = 8,75$	$\bar{Y}_{31} = 12,25$	$\bar{Y}_{.1} = 9,83$
		6	4	10	
		5	6	9	
		4	8	6	
	Medias marginales	7	5	8	
		$\bar{Y}_{12} = 5,5$	$\bar{Y}_{22} = 5,75$	$\bar{Y}_{32} = 8,25$	$\bar{Y}_{.2} = 6,5$
		$\bar{Y}_{1.} = 7$	$\bar{Y}_{2.} = 7,25$	$\bar{Y}_{3.} = 10,25$	$\bar{Y}_{..} = 8,16$

Antes de abordar el análisis de la varianza, examinaremos con más detalle la estructura de la tabla (véase la Tabla 7.2).

² Cuando hablamos de *estadío* nos referimos a la *edad en la que el niño ha aprendido su segunda lengua*. En el presente texto, hemos adaptado la clasificación que existe en función de la edad de aprendizaje de las lenguas, a fin de que el ejemplo resulte más didáctico.

TABLA 7.2 Datos correspondientes a un diseño factorial $A \times B$: modelo general

Variables independientes	a_1	...	a_j	...	a_a	$T^2_{..k}$	Puntuación media k
b_1	Y_{111} Y_{i11} Y_{n11}	...	Y_{1j1} Y_{ij1} Y_{nj1}	...	Y_{1a1} Y_{ia1} Y_{na1}		$\bar{Y}_{..1}$
...		...		...			
b_k	Y_{11k} Y_{i1k} Y_{n1k}	...	Y_{1jk} Y_{ijk} Y_{njk}	...	Y_{1ak} Y_{iak} Y_{nak}		$\bar{Y}_{..k}$
...		...		...			
b_b	Y_{11b} Y_{i1b} Y_{n1b}	...	Y_{1jb} Y_{ijb} Y_{njb}	...	Y_{1ab} Y_{iab} Y_{nab}		$\bar{Y}_{..b}$
$T^2_{.j.}$						$T^2_{...}$	
Puntuación media j	$\bar{Y}_{.1.}$		$\bar{Y}_{.j.}$		$\bar{Y}_{.a.}$		$\bar{Y}_{...}$

Cada una de las celdillas de la tabla representa *un determinado grupo*. En la parte derecha (en la columna $T^2_{..k}$) colocamos los *cuadrados de las sumas* correspondientes a las filas ($k = 1, 2, ..., b$) y en la parte inferior de la tabla (en la fila $T^2_{.j.}$), los *cuadrados de las sumas* correspondientes a las columnas ($j = 1, 2, ..., a$). Como cabe apreciar, los primeros corresponden a los *niveles del factor B* y los segundos, a las *categorías del factor A*. Con respecto a la notación, la letra *T* hace referencia a la suma de una serie de puntuaciones y, más específicamente, a la suma de los datos que se expresan mediante los subíndices asociados a dicha letra. Por ejemplo, el símbolo $T^2_{...}$ expresa *el cuadrado de la suma de todas las puntuaciones*. Por último, en los *márgenes* hemos colocado las *medias aritméticas* que corresponden a cada fila o columna.

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño, y suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 7.1.

7.3.2.3. Desarrollo del análisis factorial de la varianza

Como en el caso de los diseños unifactoriales (Epígrafes 6.1.2 y 6.2.2), calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y las razones *F* para cada uno de los parámetros de la ecuación estructural del ANOVA.

Procedimiento 1

Fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas:

- Efecto del factor A:

$$SCA = \left(\frac{1}{bn} \sum_j T_{.j}^2 \right) - C = \left[\frac{1}{bn} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C \quad (7.2)$$

- Efecto del factor B:

$$SCB = \left(\frac{1}{an} \sum_k T_{..k}^2 \right) - C = \left[\frac{1}{an} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C \quad (7.3)$$

- Efecto correspondiente a la interacción entre los factores A y B:

$$SCAB = \left[\left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C \right] - (SCA + SCB) \quad (7.4)$$

- Variabilidad residual o del error:

$$SCR = SCT - (SCA + SCB + SCAB) \quad (7.5)$$

- Variabilidad total:

$$SCT = \sum_j \sum_k \sum_i Y_{ijk}^2 - C \quad (7.6)$$

El componente C de las fórmulas precedentes se calcula mediante la siguiente expresión:

$$C = \frac{1}{N} T_{...}^2 = \frac{1}{abn} \left(\sum_j \sum_k \sum_i Y_{ijk} \right)^2 \quad (7.7)$$

Tras obtener el valor de C, llevaremos a cabo el cálculo de las sumas de cuadrados.

$$C = \frac{1}{3 \cdot 2 \cdot 4} (7 + 9 + 10 + 8 + 6 + 5 + 4 + 7 + 8 + 11 + 10 + 6 + 4 + 6 + 8 + 5 + 13 + 15 + 11 + 10 + 10 + 9 + 6 + 8)^2 = \frac{(196)^2}{24} = 1.600,67$$

Una vez obtenido este valor, procedemos a calcular la SCA, la SCB, la SCAB, la SCT y la SCR, respectivamente:

$$SCA = \frac{1}{2 \cdot 4} [(56)^2 + (58)^2 + (82)^2] - 1.600,67 = 52,33$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{.j}$, a saber, a los elementos de la fila $T_{.j}^2$ de la Tabla 7.2.

$$SCB = \frac{1}{3 \cdot 4} [(118)^2 + (78)^2] - 1.600,67 = 66,66$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{..k}$, es decir, a los elementos de la columna $T_{..k}^2$ de la Tabla 7.2.

$$SCAB = \left[\frac{1}{4} [(34)^2 + (22)^2 + (35)^2 + (23)^2 + (49)^2 + (33)^2] - 1.600,67 \right] - \\ - (52,33 + 66,66) = 1,34$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios T_{jk} , a saber, a la suma de las puntuaciones de cada celdilla de la Tabla 7.2. Como cabe observar, la cantidad de sumandos es igual a la cantidad de tratamientos de los que consta el diseño $A \times B$.

$$SCT = [(7)^2 + (9)^2 + (10)^2 + (8)^2 + (6)^2 + (5)^2 + (4)^2 + (7)^2 + (8)^2 + (11)^2 + \\ + (10)^2 + (6)^2 + (4)^2 + (6)^2 + (8)^2 + (5)^2 + (13)^2 + (15)^2 + (11)^2 + (10)^2 + \\ + (10)^2 + (9)^2 + (6)^2 + (8)^2] - 1.600,67 = 1.778 - 1.600,67 = 177,33$$

$$SCR = 177,33 - (52,33 + 66,6 + 1,34) = 57$$

Procedimiento 2: Desarrollo mediante vectores

Modificando ligeramente la secuencia empleada para el cálculo de las sumas cuadráticas en los diseños unifactoriales (Epígrafes 6.1.2 y 6.2.2), hallaremos, en primer lugar, los vectores asociados a cada uno de los parámetros de la ecuación estructural del ANOVA correspondiente al diseño factorial $A \times B$ (Fórmula 7.1) y, posteriormente, calcularemos las sumas de cuadrados de los diferentes efectos mediante los procedimientos catalogados como *procedimiento 2* y *procedimiento 3* en el caso de los diseños unifactoriales.

CÁLCULO:

VECTOR Y

Como ya es sabido, este vector no es sino el *vector columna* de todas las *puntuaciones directas*.

$$Y = \{Y\} = \begin{bmatrix} 7 \\ 9 \\ 10 \\ 8 \\ 6 \\ 5 \\ 4 \\ 7 \\ 8 \\ 11 \\ 10 \\ 6 \\ 4 \\ 6 \\ 8 \\ 5 \\ 13 \\ 15 \\ 11 \\ 10 \\ 10 \\ 9 \\ 6 \\ 8 \end{bmatrix} \quad (7.8)$$

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada una de las categorías del factor A para estimar, a partir de tales medias, los parámetros asociados al efecto del factor A , α_j .

$$\mu = \frac{1}{a \cdot b \cdot n} \sum_j \sum_k \sum_i Y_{ijk} \quad (7.9)$$

$$\mu = \left(\frac{1}{3 \cdot 2 \cdot 4} \right) [7 + 9 + 10 + 8 + 6 + 5 + 4 + 7 + 8 + 11 + 10 + 6 + \\ + 4 + 6 + 8 + 5 + 13 + 15 + 11 + 10 + 10 + 9 + 6 + 8]$$

$$\mu = \frac{196}{24} = 8,17$$

Promedios de los niveles del factor principal A :

$$\alpha_j = \mu_{j.} - \mu \quad (7.10)$$

$$\mu_{j.} = \frac{1}{b \cdot n} \sum_k \sum_i Y_{ijk}$$

respectivamente:

$$\mu_{1.} = \left(\frac{1}{2 \cdot 4} \right) [7 + 9 + 10 + 8 + 6 + 5 + 4 + 7] = 7$$

$$\mu_{2.} = \left(\frac{1}{2 \cdot 4} \right) [8 + 11 + 10 + 6 + 4 + 6 + 8 + 5] = 7,25$$

$$\mu_{3.} = \left(\frac{1}{2 \cdot 4} \right) [13 + 15 + 11 + 10 + 10 + 9 + 6 + 8] = 10,25$$

Por tanto:

$$\alpha_1 = \mu_{1.} - \mu = 7 - 8,17 = -1,17$$

$$\alpha_2 = \mu_{2.} - \mu = 7,25 - 8,17 = -0,92$$

$$\alpha_3 = \mu_{3.} - \mu = 10,25 - 8,17 = 2,08$$

El *vector* A adopta los siguientes valores:

$$A = \{\alpha\} = \begin{bmatrix} -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -1,17 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ -0,92 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \\ 2,08 \end{bmatrix} \quad (7.11)$$

VECTOR B

A fin de obtener la variabilidad correspondiente al factor B , estimaremos los parámetros asociados a dicho factor, β_k . Para ello, comenzaremos calculando los promedios de las categorías del factor B .

$$\beta_k = \mu_{.k} - \mu \quad (7.12)$$

$$\mu_{.k} = \frac{1}{a \cdot n} \sum_j \sum_i Y_{ijk}$$

respectivamente:

$$\mu_{.1} = \left(\frac{1}{3 \cdot 4} \right) [7 + 9 + 10 + 8 + 8 + 11 + 10 + 6 + 13 + 15 + 11 + 10] = 9,83$$

$$\mu_{.2} = \left(\frac{1}{3 \cdot 4} \right) [6 + 5 + 4 + 7 + 4 + 6 + 8 + 5 + 10 + 9 + 6 + 8] = 6,5$$

Por tanto:

$$\beta_1 = \mu_{.1} - \mu = 9,83 - 8,17 = 1,66$$

$$\beta_2 = \mu_{.2} - \mu = 6,5 - 8,17 = -1,67$$

El *vector* B adopta los siguientes valores:

$$B = \{\beta\} = \begin{bmatrix} 1,66 \\ 1,66 \\ 1,66 \\ 1,66 \\ -1,67 \\ -1,67 \\ -1,67 \\ -1,67 \\ 1,66 \\ 1,66 \\ 1,66 \\ 1,66 \\ -1,67 \\ -1,67 \\ -1,67 \\ -1,67 \\ 1,66 \\ 1,66 \\ 1,66 \\ 1,66 \\ -1,67 \\ -1,67 \\ -1,67 \\ -1,67 \end{bmatrix} \quad (7.13)$$

VECTOR $\{\alpha\beta\}$

La variabilidad asociada al vector $\{\alpha\beta\}$ es la correspondiente al efecto de interacción entre los factores A y B . Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\alpha\beta)_{jk}$.

$$(\alpha\beta)_{jk} = \mu_{jk} - (\mu + \alpha_j + \beta_k) \quad (7.14)$$

$$\mu_{jk} = \frac{1}{n} \sum_i Y_{ijk}$$

Promedios correspondientes a las combinaciones entre los niveles de los factores A y B :

$$\mu_{11} = \left(\frac{1}{4}\right)[7 + 9 + 10 + 8] = 8,5$$

$$\mu_{12} = \left(\frac{1}{4}\right)[6 + 5 + 4 + 7] = 5,5$$

$$\mu_{21} = \left(\frac{1}{4}\right)[8 + 11 + 10 + 6] = 8,75$$

$$\mu_{22} = \left(\frac{1}{4}\right)[4 + 6 + 8 + 5] = 5,75$$

$$\mu_{31} = \left(\frac{1}{4}\right)[13 + 15 + 11 + 10] = 12,25$$

$$\mu_{32} = \left(\frac{1}{4}\right)[10 + 9 + 6 + 8] = 8,25$$

Por tanto:

$$\begin{aligned}
 (\alpha\beta)_{11} &= \mu_{11} - \mu - \alpha_1 - \beta_1 = 8,5 - 8,17 - (-1,17) - 1,66 = -0,16 \\
 (\alpha\beta)_{12} &= \mu_{12} - \mu - \alpha_1 - \beta_2 = 5,5 - 8,17 - (-1,17) - (-1,67) = 0,17 \\
 (\alpha\beta)_{21} &= \mu_{21} - \mu - \alpha_2 - \beta_1 = 8,75 - 8,17 - (-0,92) - 1,66 = -0,16 \\
 (\alpha\beta)_{22} &= \mu_{22} - \mu - \alpha_2 - \beta_2 = 5,75 - 8,17 - (-0,92) - (-1,67) = 0,17 \\
 (\alpha\beta)_{31} &= \mu_{31} - \mu - \alpha_3 - \beta_1 = 12,25 - 8,17 - 2,08 - 1,66 = 0,34 \\
 (\alpha\beta)_{32} &= \mu_{32} - \mu - \alpha_3 - \beta_2 = 8,25 - 8,17 - 2,08 - (-1,67) = -0,33
 \end{aligned}$$

El vector $\{\alpha\beta\}$ adopta los siguientes valores:

$$\{\alpha\beta\} = \begin{bmatrix} -0,16 \\ -0,16 \\ -0,16 \\ -0,16 \\ 0,17 \\ 0,17 \\ 0,17 \\ 0,17 \\ -0,16 \\ -0,16 \\ -0,16 \\ -0,16 \\ 0,17 \\ 0,17 \\ 0,17 \\ 0,17 \\ 0,34 \\ 0,34 \\ 0,34 \\ 0,34 \\ -0,33 \\ -0,33 \\ -0,33 \\ -0,33 \end{bmatrix} \quad (7.15)$$

VECTOR E, CÁLCULO DE LOS RESIDUALES O ERRORES

Partiendo de la Fórmula (7.1):

$$\varepsilon_{ijk} = Y_{ijk} - \mu - \alpha_j - \beta_k - (\alpha\beta)_{jk}$$

$$\varepsilon_{i11} = Y_{i11} - \mu - \alpha_1 - \beta_1 - (\alpha\beta)_{11} = Y_{i11} - 8,17 - (-1,17) - 1,66 - (-0,16) = Y_{i11} - 8,5$$

Sustituyendo las puntuaciones Y_{i11} por sus correspondientes valores:

$$\varepsilon_{111} = 7 - 8,5 = -1,5$$

$$\varepsilon_{211} = 9 - 8,5 = 0,5$$

$$\varepsilon_{311} = 10 - 8,5 = 1,5$$

$$\varepsilon_{411} = 8 - 8,5 = -0,5$$

$$\varepsilon_{i12} = \mu - \alpha_1 - \beta_2 - (\alpha\beta)_{12} = Y_{i12} - 8,17 - (-1,17) - (-1,67) - 0,17 = Y_{i12} - 5,5$$

Sustituyendo las puntuaciones Y_{i12} por sus correspondientes valores:

$$\varepsilon_{112} = 6 - 5,5 = 0,5$$

$$\varepsilon_{212} = 5 - 5,5 = -0,5$$

$$\varepsilon_{312} = 4 - 5,5 = -1,5$$

$$\varepsilon_{412} = 7 - 5,5 = 1,5$$

$$\varepsilon_{i21} = Y_{i21} - \mu - \alpha_2 - \beta_1 - (\alpha\beta)_{21} = Y_{i21} - 8,17 - (-0,92) - 1,66 - (-0,16) = Y_{i21} - 8,75$$

Sustituyendo las puntuaciones Y_{i21} por sus correspondientes valores:

$$\varepsilon_{121} = 8 - 8,75 = -0,75$$

$$\varepsilon_{221} = 11 - 8,75 = 2,25$$

$$\varepsilon_{321} = 10 - 8,75 = 1,25$$

$$\varepsilon_{421} = 6 - 8,75 = -2,75$$

$$\varepsilon_{i22} = Y_{i22} - \mu - \alpha_2 - \beta_2 - (\alpha\beta)_{22} = Y_{i22} - 8,17 - (-0,92) - (-1,67) - 0,17 = Y_{i22} - 5,75$$

Sustituyendo las puntuaciones Y_{i22} por sus correspondientes valores:

$$\varepsilon_{122} = 4 - 5,75 = -1,75$$

$$\varepsilon_{222} = 6 - 5,75 = 0,25$$

$$\varepsilon_{322} = 8 - 5,75 = 2,25$$

$$\varepsilon_{422} = 5 - 5,75 = -0,75$$

$$\varepsilon_{i31} = Y_{i31} - \mu - \alpha_3 - \beta_1 - (\alpha\beta)_{31} = Y_{i31} - 8,17 - 2,08 - 1,66 - 0,34 = Y_{i31} - 12,25$$

Sustituyendo las puntuaciones Y_{i31} por sus correspondientes valores:

$$\varepsilon_{131} = 13 - 12,25 = 0,75$$

$$\varepsilon_{231} = 15 - 12,25 = 2,75$$

$$\varepsilon_{331} = 11 - 12,25 = -1,25$$

$$\varepsilon_{431} = 10 - 12,25 = -2,25$$

$$\varepsilon_{i32} = Y_{i32} - \mu - \alpha_3 - \beta_2 - (\alpha\beta)_{32} = Y_{i32} - 8,17 - 2,08 - (-1,67) - (-0,33) = Y_{i32} - 8,25$$

Sustituyendo las puntuaciones Y_{i32} por sus correspondientes valores:

$$\varepsilon_{132} = 10 - 8,25 = 1,75$$

$$\varepsilon_{232} = 9 - 8,25 = 0,75$$

$$\varepsilon_{332} = 6 - 8,25 = -2,25$$

$$\varepsilon_{432} = 8 - 8,25 = -0,25$$

Por lo tanto, el *vector E* adopta los siguientes valores:

$$E = \{\varepsilon\} = \begin{bmatrix} -1,5 \\ 0,5 \\ 1,5 \\ -0,5 \\ 0,5 \\ -0,5 \\ -1,5 \\ 1,5 \\ -0,75 \\ 2,25 \\ 1,25 \\ -2,75 \\ -1,75 \\ 0,25 \\ 2,25 \\ -0,75 \\ 0,75 \\ 2,75 \\ -1,25 \\ -2,25 \\ 1,75 \\ 0,75 \\ -2,25 \\ -0,25 \end{bmatrix} \quad (7.16)$$

Llegados a este punto, podemos calcular la *SCA*, la *SCB*, la *SCAB* y la *SCR* aplicando las Fórmulas (7.17), (7.18), (7.19) y (7.20), respectivamente, o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos.

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = bn \sum_j \alpha_j^2 = 2 \cdot 4[(-1,17)^2 + (-0,92)^2 + (2,08)^2] = 52,33 \quad (7.17)$$

donde:

$$\begin{aligned} \alpha_j &= \mu_{j.} - \mu \\ \mu &= \frac{1}{a \cdot b \cdot n} \sum_j \sum_k \sum_i Y_{ijk} \\ \mu_{j.} &= \frac{1}{b \cdot n} \sum_k \sum_i Y_{ijk} \end{aligned}$$

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR B (SCB):

$$SCB = an \sum_k \beta_k^2 = 3 \cdot 4[(1,66)^2 + (-1,67)^2] = 66,66 \quad (7.18)$$

donde:

$$\begin{aligned} \beta_k &= \mu_{.k} - \mu \\ \mu_{.k} &= \frac{1}{a \cdot n} \sum_j \sum_i Y_{ijk} \end{aligned}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y B (SCAB):

$$SCAB = n \sum_j \sum_k (\alpha\beta)_{jk}^2$$

$$SCAB = 4[(-0,16)^2 + (0,17)^2 + (-0,16)^2 + (0,17)^2 + (0,34)^2 + (-0,33)^2] = 1,34 \quad (7.19)$$

donde:

$$(\alpha\beta)_{jk} = \mu_{jk} - (\mu + \alpha_j + \beta_k)$$

$$\mu_{jk} = \frac{1}{n} \sum_i Y_{ijk}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LOS RESIDUALES O ERRORES (SCR):

$$SCR = \sum_j \sum_k \sum_i \varepsilon_{ijk}^2$$

$$SCR = (-1,5)^2 + (0,5)^2 + (1,5)^2 + \dots + (0,75)^2 + (-2,25)^2 + (-0,25)^2 = 57 \quad (7.20)$$

donde:

$$\varepsilon_{ijk} = Y_{ijk} - [\mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}]$$

No debemos olvidar que:

$$SCT = SCA + SCB + SCAB + SCR = 177,33$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza (Tabla 7.3).

TABLA 7.3 Análisis factorial de la varianza para el diseño factorial A × B: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A	SCA = 52,33	a - 1 = 2	MCA = 26,16	F _A = 8,25
Factor B	SCB = 66,66	b - 1 = 1	MCB = 66,66	F _B = 21,03
Interacción A × B	SCAB = 1,34	(a - 1)(b - 1) = 2	MCAB = 0,67	F _{AB} = 0,21
Intragrupo, residual o del error	SCR = 57	ab(n - 1) = 18	MCR = 3,17	
Total	SCT = 177,33	abn - 1 = 23		

Tras la obtención de las F observadas³, debemos determinar si la variabilidad explicada por los factores y por su interacción es o no significativa. Para ello, recurrimos a las tablas de los valores críticos de la distribución F . Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos (véase la Tabla 7.4).

TABLA 7.4 Comparaciones entre las F observadas y las F teóricas con un nivel de confianza del 95 %

Fuentes de variación	F crítica _(0,95; gl1/gl2)	F observada	Diferencia
Factor A	$F_{0,95;2/18} = 3,55$	$F_A = 8,25$	$8,25 > 3,55$
Factor B	$F_{0,95;1/18} = 4,41$	$F_B = 21,03$	$21,03 > 4,41$
Interacción A $\times$ B	$F_{0,95;2/18} = 3,55$	$F_{AB} = 0,21$	$0,21 < 3,55$

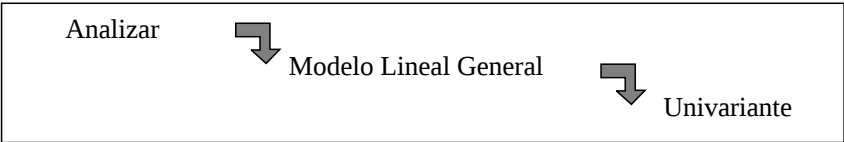
Como puede observarse en la Tabla 7.4, tanto el nivel de competencia de los padres en la segunda lengua (factor A) como el estadio lingüístico (factor B), ejercen una influencia estadísticamente significativa ($p \leq 0,05$) sobre la competencia que presenta el niño en su segunda lengua. No obstante, no se observa un efecto de interacción entre ambos factores. Teniendo en cuenta que la variable nivel de competencia de los padres tiene tres niveles, para concluir adecuadamente la interpretación de los resultados deberíamos utilizar algún procedimiento de comparaciones múltiples (véase el Epígrafe 6.2.2.3), tal y como se ilustra en el siguiente apartado.

7.3.2.4. *Análisis de datos mediante el paquete estadístico SPSS 10.0*

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

Análisis factorial de la varianza

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



³ Cabe advertir que el hecho de aplicar la técnica del redondeo de decimales en el cálculo de los diferentes componentes de los análisis de la varianza que se desarrollan a lo largo del texto, puede generar pequeñas diferencias en los decimales de las F observadas que se presentan en las tablas elaboradas manualmente y en las obtenidas mediante el programa estadístico SPSS 10.0.

- Indicamos la(s) variable(s) dependiente(s) y los factores, especificando en su lugar correspondiente si son fijos o aleatorios (en nuestro ejemplo, ambos factores son fijos).
- El menú **Gráficos** nos permite representar gráficamente los resultados de la interacción entre los factores. Para ello, debemos escoger la variable que deseamos que esté representada en el *eje horizontal* (en nuestro ejemplo, el factor «Nivel de competencia de los padres»), así como la variable que queremos representar mediante *líneas distintas* (en nuestro ejemplo, el factor «Tipología lingüística»). A continuación seleccionamos la opción *Añadir* y posteriormente la opción *Continuar*.



- El menú **Post-hoc** incluye distintos procedimientos de comparaciones múltiples. No obstante, cabe señalar que dichos procedimientos establecen comparaciones entre los niveles de los diferentes factores por separado, ya que el programa SPSS 10.0 no proporciona la posibilidad de realizar comparaciones múltiples entre los tratamientos que resultan de la combinación entre dichos factores. En el caso de que la interacción resultara estadísticamente significativa, con el fin de realizar todas las comparaciones múltiples dos a dos, deberíamos llevar a cabo una transformación en las variables. Es decir, deberíamos crear una nueva variable cuyas categorías correspondiesen a las distintas combinaciones entre los niveles de los factores. Dicho procedimiento se expondrá al final del presente epígrafe.
- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor así como para la interacción. Asimismo, permite seleccionar de entre un conjunto de opciones aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar* «estadísticos descriptivos», «estimaciones del tamaño del efecto», «potencia observada» y «pruebas de homogeneidad»).



- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

```
UNIANOVA
competencia BY padres tipolog
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PLOT = PROFILE( padres*tipolog )
/EMMEANS = TABLES(padres)
/EMMEANS = TABLES(tipolog)
/PRINT = DESCRIPTIVE ETASQ OPOWER HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = padres tipolog padres*tipolog .
```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Nivel de competencia de los padres	1	Sólo es competente la madre	8
	2	Sólo es competente el padre	8
	3	Ambos progenitores son competentes	8
Tipología lingüística	1	Temprano	12
	2	Tardío	12

Estadísticos descriptivos**Variable dependiente: Nivel de competencia en la segunda lengua**

Nivel de competencia de los padres	Tipología lingüística	Media	Desv. típ.	N
Sólo es competente la madre	Precoz	8,50	1,29	4
	Tardío	5,50	1,29	4
	Total	7,00	2,00	8
Sólo es competente el padre	Precoz	8,75	2,22	4
	Tardío	5,75	1,71	4
	Total	7,25	2,43	8
Ambos progenitores son competentes	Precoz	12,25	2,22	4
	Tardío	8,25	1,71	4
	Total	10,25	2,82	8
Total	Precoz	9,83	2,52	12
	Tardío	6,50	1,93	12
	Total	8,17	2,78	24

Contraste de Levene sobre la igualdad de las varianzas error^a**Variable dependiente: Nivel de competencia en la segunda lengua**

F	gl1	gl2	Sig.
0,700	5	18	0,631

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a lo largo de todos los grupos.

^a Diseño: Intercept + PADRES + TIPOLOG + PADRES * TIPOLOG.

Pruebas de los efectos intersujetos**Variable dependiente: Nivel de competencia en la segunda lengua**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	120,333 ^b	5	24,067	7,600	0,001	0,679	38,000	0,992
Intercept	1.600,667	1	1.600,667	505,474	0,000	0,966	505,474	1,000
PADRES	52,333	2	26,167	8,263	0,003	0,479	16,526	0,926
TIPOLOG	66,667	1	66,667	21,053	0,000	0,539	21,053	0,991
PADRES * TIPOLOG	1,333	2	0,667	0,211	0,812	0,023	0,421	0,078
Error	57,000	18	3,167					
Total	1.778,000	24						
Total corregido	177,333	23						

^a Calculado con alfa = 0,05.

^b R cuadrado = 0,679 (R cuadrado corregido = 0,589).

Medias marginales estimadas

1. Nivel de competencia de los padres

Variable dependiente: Nivel de competencia en la segunda lengua

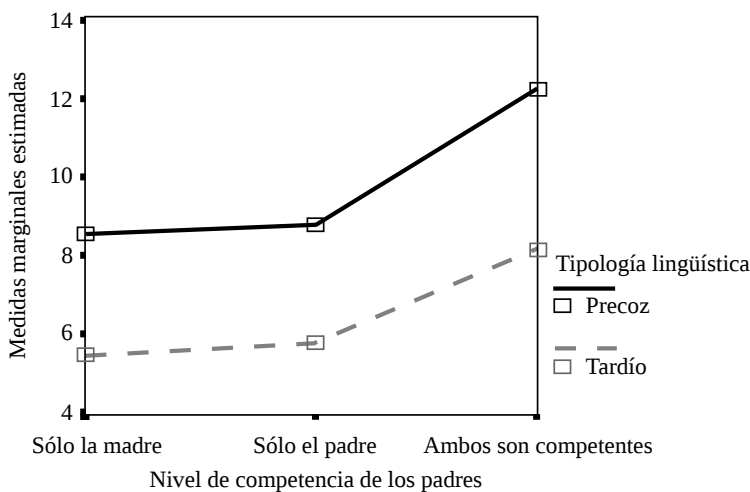
Nivel de competencia de los padres	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Sólo es competente la madre	7,000	0,629	5,678	8,322
Sólo es competente el padre	7,250	0,629	5,928	8,572
Ambos progenitores son competentes	10,250	0,629	8,928	11,572

2. Tipología lingüística

Variable dependiente: Nivel de competencia en la segunda lengua

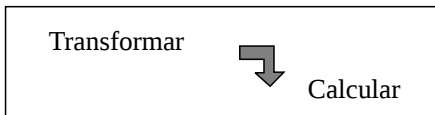
Tipología lingüística	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Precoz	9,833	0,514	8,754	10,913
Tardío	6,500	0,514	5,421	7,579

Gráficos de perfil



Comparaciones múltiples

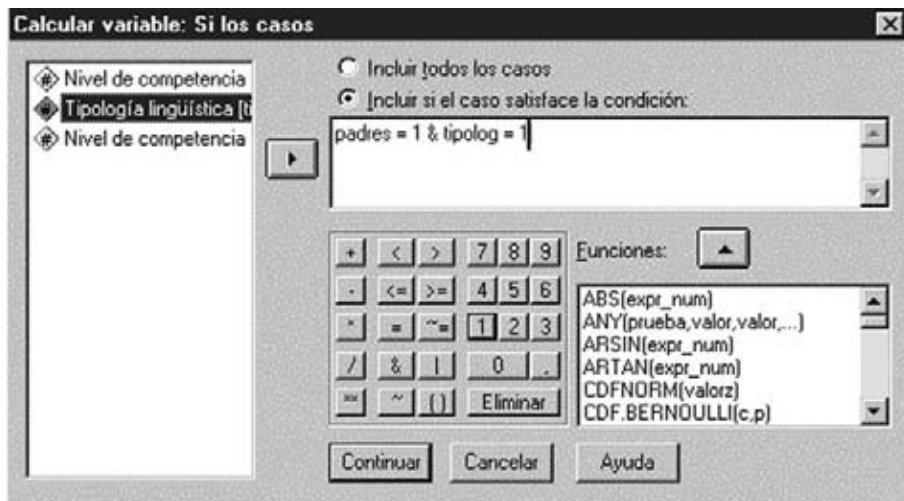
- Observando los resultados obtenidos, constatamos que existen dos efectos principales significativos, correspondientes a los dos factores incluidos en el diseño. Por el contrario, la interacción entre ambos factores no resulta estadísticamente significativa. Dado que el factor «tipología lingüística» consta de dos niveles, únicamente deben realizarse pruebas de comparaciones múltiples para los niveles del factor «nivel de competencia de los padres». Para ello, han de seguirse las pautas descritas en el capítulo anterior, tanto en lo que respecta a la elección del procedimiento (Epígrafe 6.2.2.4) como a su cálculo mediante el paquete estadístico SPSS 10.0 (Epígrafe 6.2.2.5).
- Como ya se ha indicado anteriormente, en el caso de que la interacción resultara estadísticamente significativa y de que estuviéramos interesados en realizar todas las posibles comparaciones dos a dos, deberíamos llevar a cabo una transformación en las variables, con el fin de crear una nueva variable cuyos niveles correspondiesen a los tratamientos resultantes de la combinación entre los niveles de los factores A y B. Dicho de otro modo, la combinación entre los niveles del factor A y del factor B dará lugar a los niveles de la nueva variable que se crea mediante el programa SPSS. Para llevar a cabo dicha transformación, escogemos las opciones **Transformar** y **Calcular**, sucesivamente.



- En el cuadro de diálogo que aparece en pantalla, debemos darle un nombre a la nueva *Variable de destino* (en nuestro ejemplo, la llamaremos «interacción»).
- A continuación, en el cuadro que lleva por título *Expresión numérica*, asignamos un determinado valor a dicha variable (en nuestro ejemplo, el primer valor será 1).



- Seguidamente, escogemos la opción Si...



- En este cuadro de diálogo, escogemos la opción *Incluir si el caso satisface la condición*. A continuación, debemos establecer la condición que representa al valor 1, en el cuadro correspondiente (en nuestro ejemplo, la primera condición será Padres = 1 & Tipología = 1). Seguidamente, escogemos la opción *Continuar*, que nos lleva al primer cuadro de diálogo. En éste, seleccionamos la opción *Aceptar*.

Mediante todas estas operaciones, hemos creado una nueva variable denominada *Interacción*, cuyo valor 1 representa la condición «Sólo la madre es competente» (Padres = 1) y «Nivel de adquisición precoz» (Tipología = 1).

Para concluir la creación de la nueva variable, y de sus correspondientes categorías, debemos repetir estos pasos con el resto de los niveles, de modo que la nueva variable *Interacción* tenga 6 niveles. En nuestro ejemplo, la sintaxis que nos permite crear dicha variable es la siguiente:

```
IF (padres = 1 & tipolog = 1) interac = 1.
EXECUTE.
IF (padres = 1 & tipolog = 2) interac = 2.
EXECUTE.
IF (padres = 2 & tipolog = 1) interac = 3.
EXECUTE.
IF (padres = 2 & tipolog = 2) interac = 4.
EXECUTE.
IF (padres = 3 & tipolog = 1) interac = 5.
EXECUTE.
IF (padres = 3 & tipolog = 2) interac = 6.
EXECUTE.
```

- Finalizado el proceso se pueden realizar las comparaciones múltiples mediante el procedimiento pertinente, tal y como se ha expuesto en el Epígrafe 6.2.2.5.

7.3.3. Análisis factorial de la varianza para el diseño factorial $A \times B \times C$

7.3.3.1. Modelo general de análisis

Cuando el diseño consta de tres factores, el modelo matemático subyacente a la predicción que se realiza bajo la hipótesis alternativa puede representarse mediante la siguiente ecuación:

$$y_{ijkm} = \mu + \alpha_j + \beta_k + \gamma_m + (\alpha\beta)_{jk} + (\alpha\gamma)_{jm} + (\beta\gamma)_{km} + (\alpha\beta\gamma)_{jkm} + \varepsilon_{ijkm} \quad (7.21)$$

donde:

y_{ijkm} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable A , el k -ésimo nivel de la variable B y el m -ésimo nivel de la variable C .

μ = Media general de la variable dependiente.

α_j = Efecto principal asociado a la administración del j -ésimo nivel de la variable A .

β_k = Efecto principal asociado a la administración del k -ésimo nivel de la variable B .

γ_m = Efecto principal asociado a la administración del m -ésimo nivel de la variable C .

$(\alpha\beta)_{jk}$ = Efecto producido por la interacción entre el j -ésimo nivel de A y el k -ésimo nivel de B .

$(\alpha\gamma)_{jm}$ = Efecto producido por la interacción entre el j -ésimo nivel de A y el m -ésimo nivel de C .

$(\beta\gamma)_{km}$ = Efecto producido por la interacción entre el k -ésimo nivel de B y el m -ésimo nivel de C .

$(\alpha\beta\gamma)_{jkm}$ = Efecto producido por la interacción entre el j -ésimo nivel de A , el k -ésimo nivel de B y el m -ésimo nivel de C .

ε_{ijkm} = Término de error o componente aleatorio del modelo. Se asume que $\varepsilon_{ijkm} \simeq NID(O, \sigma_e^2)$.

La lógica es similar a la del diseño factorial de dos factores. Así, partiendo del *modelo completo*, la variación total se descompone en dos componentes: la *variación intergrupos* y la *variación intragrupo*. Sin embargo, en el diseño factorial de tres factores, la variación intergrupos incluye muchas más fuentes de variación que en el diseño de dos factores. En concreto, incluye siete fuentes de variación: las variaciones debidas a cada uno de los factores (A , B y C), las debidas a cada una de las interacciones de primer orden ($A \times B$, $A \times C$ y $B \times C$) y la derivada de la interacción de segundo orden ($A \times B \times C$). La variación intragrupo, por su parte, incluye únicamente la suma cuadrática del componente residual del modelo. A partir de las sumas de cuadrados y de sus correspondientes grados de libertad se obtienen las medias cuadráticas de cada uno de los parámetros del modelo y se dividen tales varianzas entre la media cuadrática del error. En el análisis de este diseño factorial se comienza comprobando si la interacción de orden superior resulta significativa. En caso afirmativo, cualquier interpretación acerca de las interacciones de orden inferior y de los efectos principales queda mediatizada por la presencia de dicha interacción. Si la interacción de orden superior no resulta significativa se deben contrastar las hipótesis de nulidad referidas a las interacciones de orden inferior. Cuando se obtienen resultados estadísticamente significativos en esta fase del análisis, los efectos principales no deben tenerse en cuenta. Por el contrario, si tales interacciones no resultan estadísticamente significativas, se deben interpretar los efectos de cada uno de los factores de manera independiente.

7.3.3.2. Ejemplo práctico

Supongamos que en el ámbito de la psicología clínica, realizamos una investigación para examinar la influencia que ejercen *tres tipos de terapias* distintas, sobre el *estado físico y psicológico de sujetos que padecen SIDA (Síndrome de Inmuno-Deficiencia Adquirida)*. Para ello se asignan, aleatoriamente, la mitad de los sujetos de la muestra a un grupo que *no recibe ningún tipo de fármaco* (a_1), mientras que la otra mitad *se somete a una terapia basada en fármacos* (a_2). A su vez, cada uno de estos grupos se subdivide en dos subgrupos en función de las categorías que adopta el *factor B* o la *terapia individual*. Así, la mitad de los enfermos de cada una de las condiciones anteriores se asigna al azar a un grupo que *recibe terapia individual* (b_2), mientras que el resto *no recibe dicha terapia* (b_1). Por último, se manipula un tercer factor (*factor C*) consistente en aplicar una *terapia familiar*. De esta forma, se vuelven a subdividir aleatoriamente los grupos en función de los dos niveles que adopta este factor, a saber: (c_1) *ausencia de terapia familiar* y (c_2) *aplicación de terapia familiar*. Siguiendo los criterios citados, se forman 8 grupos experimentales de 3 sujetos cada uno. Tras la aplicación de los tratamientos, se mide el estado físico y psicológico de los enfermos (variable criterio). En la Tabla 7.5 pueden observarse los resultados obtenidos en la investigación.

La Tabla 7.5 refleja la estructura que corresponde a este modelo de diseño. En las filas se representan las *categorías del factor C* (c_1 y c_2) y en las columnas las de los otros dos factores que configuran el diseño: en la parte superior los *dos niveles del factor A* (a_1 y a_2) y debajo de éstos, los *dos niveles del factor B* (b_1 y b_2). Por tanto, los subíndices correspondientes a los factores A, B y C son, $j = 1, 2$; $k = 1, 2$ y $m = 1, 2$, respectivamente.

TABLA 7.5 Matriz de datos del experimento

		A (Terapia basada en fármacos)				Medias marginales
		a_1 (No recibe terapia)		a_2 (Recibe terapia)		
B Terapia individual		b_1 (No)	b_2 (Sí)	b_1 (No)	b_2 (Sí)	
C Terapia familiar	c_1 (No)	22	14	16	28	$\bar{Y}_{..1} = 19,25$
		20	13	18	29	
		18	14	14	25	
		$\bar{Y}_{111} = 20$	$\bar{Y}_{121} = 13,66$	$\bar{Y}_{211} = 16$	$\bar{Y}_{221} = 27,33$	
	c_2 (Sí)	9	35	34	12	$\bar{Y}_{..2} = 21,91$
		10	29	30	16	
		8	33	32	15	
		$\bar{Y}_{112} = 9$	$\bar{Y}_{122} = 32,33$	$\bar{Y}_{212} = 32$	$\bar{Y}_{222} = 14,33$	
Medias marginales		$\bar{Y}_{11.} = 14,5$	$\bar{Y}_{12.} = 23$	$\bar{Y}_{21.} = 24$	$\bar{Y}_{22.} = 20,83$	$\bar{Y}_{...} = 20,58$

La Tabla 7.6 permite vislumbrar, con mayor claridad, la estructura que subyace al tipo de diseño que nos ocupa.

Resumiendo lo que expresan los subíndices:

$i = 1, 2, 3, \dots, n$ (sujetos dentro de cada tratamiento experimental).

$j = 1, 2, 3, \dots, a$ (niveles del factor A).

$k = 1, 2, 3, \dots, b$ (niveles del factor B).

$m = 1, 2, 3, \dots, c$ (niveles del factor C).

En la Tabla 7.6 puede apreciarse la complejidad que se deriva de las dimensiones del diseño.

A fin de facilitar la comprensión de la notación y de las fórmulas que se utilizan en el desarrollo del análisis de la varianza para este tipo de diseño, presentamos la Tabla 7.7 en la que se puede observar cómo se calculan los sumatorios marginales y los cuadrados de tales sumatorios.

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño, y suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 7.5.

TABLA 7.6 Datos correspondientes a un diseño factorial $A \times B \times C$: modelo general

Factor A									
a_1			a_j			a_a			
Factor B	b_1	b_k	b_b	b_1	b_k	b_b	b_1	b_k	b_b
Factor C									
c_1	Y_{1111}	Y_{11k1}	Y_{11b1}	Y_{1j11}	Y_{1jk1}	Y_{1jb1}	Y_{1a11}	Y_{1ak1}	Y_{1ab1}
	Y_{i111}	Y_{i1k1}	Y_{i1b1}	Y_{ij11}	Y_{ijk1}	Y_{ijb1}	Y_{ia11}	Y_{iak1}	Y_{iab1}
	Y_{n111}	Y_{n1k1}	Y_{n1b1}	Y_{nj11}	Y_{nj11}	Y_{njb1}	Y_{na11}	Y_{nak1}	Y_{nab1}
c_m	Y_{111m}	Y_{11km}	Y_{11bm}	Y_{1j1m}	Y_{1jkm}	Y_{1jbm}	Y_{1a1m}	Y_{1akm}	Y_{1abm}
	Y_{i11m}	Y_{i1km}	Y_{i1bm}	Y_{ij1m}	Y_{ijkm}	Y_{ijbm}	Y_{ia1m}	Y_{iakm}	Y_{iabm}
	Y_{n11m}	Y_{n1km}	Y_{n1bm}	Y_{nj1m}	Y_{njkm}	Y_{njbm}	Y_{na1m}	Y_{nakm}	Y_{nabm}
c_c	Y_{111c}	Y_{11kc}	Y_{11bc}	Y_{1j1c}	Y_{1jkc}	Y_{1jbc}	Y_{1a1c}	Y_{1akc}	Y_{1abc}
	Y_{i11c}	Y_{i1kc}	Y_{i1bc}	Y_{ij1c}	Y_{ijkc}	Y_{ijbc}	Y_{ia1c}	Y_{iakc}	Y_{iabc}
	Y_{n11c}	Y_{n1kc}	Y_{n1bc}	Y_{nj1c}	Y_{njkc}	Y_{njb1}	Y_{na1c}	Y_{nake}	Y_{nabc}

TABLA 7.7 Sumatorios y cuadrados de sumatorios marginales de los datos correspondientes a un diseño factorial $A \times B \times C$

	A									$T^2_{\dots m}$
	a_1			a_j			a_a			
B	b_1	b_k	b_b	b_1	b_k	b_b	b_1	b_k	b_b	
C										
c_1										$T^2_{\dots 1}$
c_m										$T^2_{\dots m}$
c_c										$T^2_{\dots c}$
$T^2_{.j..}$	$T^2_{.1..}$			$T^2_{.j..}$			$T^2_{.a..}$			
$T^2_{..k.}$	$T_{.11.}$			$T_{.j1.}$			$T_{.a1.}$			$T^2_{..1.}$
		$T_{.1k.}$			$T_{.jk.}$			$T_{.ak.}$		$T^2_{..k.}$
			$T_{.1b.}$			$T_{.jb.}$			$T_{.ab.}$	$T^2_{..b.}$
Total										$T^2_{\dots}$

7.3.3.3. Desarrollo del análisis factorial de la varianza

En el presente diseño, seguiremos la misma secuencia empleada en el caso del diseño factorial $A \times B$ para el cálculo de las sumas de cuadrados, de las varianzas y de las razones F asociadas a los distintos parámetros de la ecuación estructural del ANOVA. No obstante, realizaremos algunos comentarios adicionales acerca del efecto de interacción y de la importancia de su adecuada fundamentación teórica.

Procedimiento 1

Fórmulas necesarias para el cálculo de las sumas cuadráticas:

- Efecto del factor A:

$$SCA = \left(\frac{1}{bcn} \sum_j T^2_{.j..} \right) - C = \left[\frac{1}{bcn} \sum_j \left(\sum_i \sum_k \sum_m Y_{ijkm} \right)^2 \right] - C \quad (7.22)$$

- Efecto del factor B:

$$SCB = \left(\frac{1}{acn} \sum_k T^2_{..k.} \right) - C = \left[\frac{1}{acn} \sum_k \left(\sum_i \sum_j \sum_m Y_{ijkm} \right)^2 \right] - C \quad (7.23)$$

- Efecto del factor C:

$$SCC = \left(\frac{1}{abn} \sum_m T_{...m}^2 \right) - C = \left[\frac{1}{abn} \sum_m \left(\sum_i \sum_j \sum_k Y_{ijkm} \right)^2 \right] - C \quad (7.24)$$

- Efecto correspondiente a la interacción entre los factores A y B:

$$SCAB = \left(\left[\frac{1}{cn} \sum_j \sum_k \left(\sum_i \sum_m Y_{ijkm} \right)^2 \right] - C \right) - (SCA + SCB) \quad (7.25)$$

- Efecto correspondiente a la interacción entre los factores A y C:

$$SCAC = \left(\left[\frac{1}{bn} \sum_j \sum_m \left(\sum_i \sum_k Y_{ijkm} \right)^2 \right] - C \right) - (SCA + SCC) \quad (7.26)$$

- Efecto correspondiente a la interacción entre los factores B y C:

$$SCBC = \left(\left[\frac{1}{an} \sum_k \sum_m \left(\sum_i \sum_j Y_{ijkm} \right)^2 \right] - C \right) - (SCB + SCC) \quad (7.27)$$

- Efecto correspondiente a la interacción entre los factores A, B y C:

$$SCABC = \left(\left[\frac{1}{n} \sum_j \sum_k \sum_m \left(\sum_i Y_{ijkm} \right)^2 \right] - C \right) - (SCA + SCB + SCC + SCAB + SCAC + SCBC) \quad (7.28)$$

- Variabilidad residual o del error:

$$SCR = \sum_j \sum_k \sum_m \sum_i Y_{ijkm}^2 - \frac{1}{n} \sum_j \sum_k \sum_m \left(\sum_i Y_{ijkm} \right)^2 \quad (7.29)$$

$$SCR = SCT - (SCA + SCB + SCC + SCAB + SCAC + SCBC + SCABC)$$

- Variabilidad total:

$$SCT = \sum_j \sum_k \sum_m \sum_i Y_{ijkm}^2 - C \quad (7.30)$$

El componente C de las fórmulas precedentes se calcula mediante la siguiente expresión:

$$C = \frac{1}{N} T^2_{....} = \frac{1}{abcn} \left(\sum_j \sum_k \sum_m \sum_i Y_{ijkm} \right)^2 \quad (7.31)$$

Tras obtener el valor de C , llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{2 \cdot 2 \cdot 2 \cdot 3} (22 + 20 + 18 + 9 + 10 + 8 + 14 + 13 + 14 + 35 + 29 + 33 + 16 + 18 + 14 + 34 + 30 + 32 + 28 + 29 + 25 + 12 + 16 + 15)^2 = \frac{(494)^2}{24} = 10.168,167$$

Una vez obtenido este valor, procedemos a calcular la SCA , la SCB y la SCC , respectivamente:

$$SCA = \frac{1}{bcn} [T^2_{.1..} + T^2_{.2..}] - C$$

$$SCA = \frac{1}{2 \cdot 2 \cdot 3} [(225)^2 + (269)^2] - 10.168,167 = 80,66$$

Para saber a qué corresponden los elementos $T^2_{.j..}$ y $T^2_{..k.}$, puede recurrirse a la Tabla 7.7 o bien a la Tabla 7.8 que se presenta al final de este epígrafe. Dando por hecho que tal notación resulta ya familiar para el lector, calcularemos la variabilidad correspondiente al factor B .

$$SCB = \frac{1}{acn} [T^2_{..1.} + T^2_{..2.}] - C$$

$$SCB = \frac{1}{2 \cdot 2 \cdot 3} [(231)^2 + (263)^2] - 10.168,167 = 42,66$$

Como ya es sabido, estamos calculando los componentes aditivos de la variabilidad total. El siguiente cálculo corresponde a la variabilidad asociada al factor C . Procedamos a su estimación:

$$SCC = \frac{1}{abn} [T^2_{...1} + T^2_{...2}] - C$$

$$SCC = \frac{1}{2 \cdot 2 \cdot 3} [(231)^2 + (263)^2] - 10.168,167 = 42,66$$

Aunque en el presente ejemplo hemos obtenido valores idénticos en la SCB y en la SCC , ello no significa que siempre deba darse esta circunstancia.

A continuación, procederemos al cálculo de las sumas cuadráticas correspondientes a las interacciones entre los factores, a saber, $SCAB$, $SCAC$, $SCBC$ y $SCABC$.

Interacción entre los factores A y B:

$$SCAB = \left[\frac{1}{cn} (T_{..11}^2 + T_{..12}^2 + T_{..21}^2 + T_{..22}^2) \right] - (C + SCA + SCB)$$

$$SCAB = \left[\frac{1}{2 \cdot 3} ((87)^2 + (138)^2 + (144)^2 + (125)^2) \right] - (10.168,167 + 80,66 + 42,66) = 204,17$$

Interacción entre los factores A y C:

$$SCAC = \left[\frac{1}{bn} (T_{.1.1}^2 + T_{.1.2}^2 + T_{.2.1}^2 + T_{.2.2}^2) \right] - (C + SCA + SCC)$$

$$SCAC = \left[\frac{1}{2 \cdot 3} ((101)^2 + (124)^2 + (130)^2 + (139)^2) \right] - (10.168,167 + 80,66 + 42,66) = 8,17$$

Interacción entre los factores B y C:

$$SCBC = \left[\frac{1}{an} (T_{..11}^2 + T_{..12}^2 + T_{..21}^2 + T_{..22}^2) \right] - (C + SCB + SCC)$$

$$SCBC = \left[\frac{1}{2 \cdot 3} ((108)^2 + (123)^2 + (123)^2 + (140)^2) \right] - (10.168,167 + 42,66 + 42,66) = 0,17$$

Interacción entre los factores A, B y C:

$$SCABC = \left[\frac{1}{n} (T_{.111}^2 + T_{.112}^2 + T_{.121}^2 + T_{.122}^2 + T_{.211}^2 + T_{.212}^2 + T_{.221}^2 + T_{.222}^2) \right] -$$

$$- (C + SCA + SCB + SCC + SCAB + SCAC + SCBC)$$

$$SCABC = \left[\frac{1}{3} ((60)^2 + (27)^2 + (41)^2 + (97)^2 + (48)^2 + (96)^2 + (82)^2 + (43)^2) \right] -$$

$$- (10.168,167 + 80,66 + 42,66 + 42,66 + 204,17 + 8,17 + 0,17)$$

$$SCABC = 11.837,333 - 10.546,657 = 1.290,66$$

Como cabe observar, la cantidad de sumandos T es igual a la cantidad de condiciones o grupos experimentales de los que consta el diseño $A \times B \times C$. En nuestro caso, dicho resultado se obtiene multiplicando los niveles de los diferentes factores, a saber, $2 \times 2 \times 2 = 8$.

Seguidamente, procederemos al cálculo de la variabilidad residual o del error (SCR). Partiendo de la Fórmula (7.29), calcularemos ambos componentes de la resta. Para ello, comenzaremos con la *suma de los cuadrados* de las puntuaciones:

$$\sum_j \sum_k \sum_m \sum_i Y_{ijkm}^2 = [(22)^2 + (20)^2 + (18)^2 + (9)^2 + (10)^2 + (8)^2 + (14)^2 + (13)^2 +$$

$$+ (14)^2 + (35)^2 + (29)^2 + (33)^2 + (16)^2 + (18)^2 + (14)^2 +$$

$$+ (34)^2 + (30)^2 + (32)^2 + (28)^2 + (29)^2 + (25)^2 + (12)^2 +$$

$$+ (16)^2 + (15)^2] = 11.900$$

Con respecto al segundo componente de la citada resta, cabe señalar que es igual al primer componente de la Fórmula (7.28) utilizada para calcular la variabilidad asociada a la interacción $A \times B \times C$:

$$\frac{1}{n} \sum_j \sum_k \sum_m \left(\sum_i Y_{ijk} \right)^2 = 11.837,333$$

Una vez calculados ambos componentes, la variabilidad correspondiente a los errores (*SCR*) es:

$$SCR = 11.900 - 11.837,333 = 62,67$$

Por último, procederemos al cálculo de la variabilidad total. Ya sabemos que la variabilidad total (*SCT*) puede obtenerse mediante la suma de las variabilidades asociadas al resto de los efectos,

$$SCT = 80,66 + 42,66 + 42,66 + 204,17 + 8,17 + 0,17 + 1.290,66 + 62,67 = 1.731,83$$

o bien aplicando la Fórmula (7.30).

$$SCT = 11.900 - 10.168,167 = 1.731,83$$

Para finalizar con el cálculo de las sumas de cuadrados mediante este procedimiento, nos ha parecido interesante exponer en una tabla (véase la Tabla 7.8) cómo se obtienen los *sumatorios T* y los *cuadrados de tales sumatorios*. Aunque la tabla refleja la estructura

TABLA 7.8 Sumatorios y cuadrados de sumatorios marginales de los datos correspondientes a un diseño factorial $2 \times 2 \times 2$

	A				$T^2_{...m}$
	a_1		a_2		
	b_1	b_2	b_1	b_2	
B	b_1	b_2	b_1	b_2	
C					
c_1	$T_{.111}$	$T_{.121}$	$T_{.211}$	$T_{.221}$	$T^2_{...1}$
c_2	$T_{.112}$	$T_{.122}$	$T_{.212}$	$T_{.222}$	$T^2_{...2}$
$T^2_{.j..}$	$T^2_{.1..}$		$T^2_{.2..}$		
$T^2_{..k.}$	$T_{.11.}$		$T_{.21.}$		$T^2_{..1.}$
		$T_{.12.}$		$T_{.22.}$	$T^2_{..2.}$
Total					$T^2_{....}$

correspondiente al modelo de diseño de nuestro ejemplo práctico, consideramos que puede ser fácilmente generalizable a cualquier otro tipo de diseño.

Como cabe observar, a medida que aumenta la complejidad del modelo de diseño, las tablas también resultan más complicadas. De cualquier forma, la estructuración adecuada de los datos facilita enormemente los cálculos, sobre todo cuando éstos se hacen de forma manual.

Por tanto, resulta muy recomendable organizar adecuadamente las tablas de datos antes de proceder a su análisis. Incluso en el supuesto de que utilicemos el ordenador para dicho análisis, no debemos olvidar que la presentación adecuada de los datos facilita en gran medida la comprensión de cualquier informe de investigación.

Aun a riesgo de ser reiterativos, recordemos que la *variabilidad total* incluye la *variabilidad asociada a los efectos de los factores principales*, la *variabilidad asociada a las interacciones entre los factores* y la *variabilidad debida a los errores*. Aunque no siempre se producen *interacciones entre los factores*, nosotros hemos decidido desarrollar el ANOVA a partir de los *modelos no aditivos* o que tienen en cuenta tales interacciones. No obstante, queremos señalar que el hecho de que se produzcan o no interacciones no es una mera cuestión estadística, sino que debe entenderse, básicamente, como una consecuencia derivada del modelo teórico planteado por el investigador. Así, teniendo en cuenta el tipo de diseño y la naturaleza de los factores con los que trabaja, el investigador debe predecir, *a priori*, si van a darse o no interacciones entre tales factores.

Supongamos, por ejemplo, que realizamos una investigación para examinar la posible influencia de los factores *tipología del hablante* (monolingüe, bilingüe simultáneo, bilingüe precoz y bilingüe tardío), *sexo* (hombre, mujer) y *lugar de nacimiento* (País Vasco u otro lugar), sobre la *competencia que presenta el sujeto en euskera*⁴. Aunque cada uno de estos factores puede tener algún efecto sobre la variable criterio, no cabe suponer a priori que existe interacción entre todos ellos. Es evidente que, entre la *tipología del hablante* y el *lugar de nacimiento*, existe interacción. Sin embargo, es prácticamente impensable que se produzca algún tipo de interacción entre el *lugar de nacimiento* y el *sexo*. De hecho, partiendo de una base teórica, tales variables son independientes entre sí y es el criterio teórico el que debe utilizarse, principalmente, para establecer, a priori, si se va a producir interacción entre distintas variables. El hecho de observar que una interacción es estadísticamente significativa no permite asegurar que ésta se produce realmente, sobre todo cuando existen razones teóricas para pensar que los factores implicados en la misma son independientes entre sí. En consecuencia, la interacción debe tenerse en cuenta cuando, además de resultar significativa desde el punto de vista estadístico, es teóricamente aceptable. Retomando el ejemplo arriba propuesto, la existencia de una interacción entre la *tipología del hablante* y el *lugar de nacimiento* expresaría que la influencia que ejerce cualquiera de tales factores sobre la *competencia que presenta el sujeto en euskera* cambia en función de los valores que adopta el otro factor. Es decir, que el efecto de cada factor varía en la medida en que se modifican los niveles del otro factor y que, por tanto, no podemos afirmar que la *tipología del hablante* o el *lugar de nacimiento* ejerzan por sí solos algún tipo de efecto sobre la variable criterio. En definitiva, la interacción no representa la mera adición de los efectos que ejercen los factores por separado, sino que aporta una nueva información que puede resultar muy útil para explicar la variabilidad de la variable objeto de estudio. Dejando un análisis más exhaustivo de la interacción, desde el punto de vista estadístico,

⁴ El euskera es, junto con el español, una de las dos lenguas oficiales del País Vasco.

VECTOR B

A fin de obtener la variabilidad correspondiente al factor B , estimaremos los parámetros asociados a dicho factor, β_k . Para ello, comenzaremos calculando los promedios de los niveles del factor principal B , que son:

$$\mu_{.1.} = \left(\frac{1}{2 \cdot 2 \cdot 3} \right) [22 + 20 + 18 + 9 + 10 + 8 + 16 + 18 + 14 + 34 + 30 + 32] = 19,25$$

$$\mu_{.2.} = \left(\frac{1}{2 \cdot 2 \cdot 3} \right) [14 + 13 + 14 + 35 + 29 + 33 + 28 + 29 + 25 + 12 + 16 + 15] = 21,91$$

Por tanto:

$$\beta_1 = \mu_{.1.} - \mu = 19,25 - 20,58 = -1,33$$

$$\beta_2 = \mu_{.2.} - \mu = 21,91 - 20,58 = 1,33$$

El *vector B* adopta los siguientes valores:

$$B = \{\beta\} = \begin{bmatrix} -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ 1,33 \end{bmatrix} \quad (7.36)$$

VECTOR C

A fin de obtener la variabilidad correspondiente al factor C , estimaremos los parámetros asociados a dicho factor, γ_m . Como en los casos anteriores, calculamos, en primer lugar, los promedios de las categorías del factor principal C :

$$\mu_{..1} = \left(\frac{1}{2 \cdot 2 \cdot 3} \right) [22 + 20 + 18 + 14 + 13 + 14 + 16 + 18 + 14 + 28 + 29 + 25] = 19,25$$

$$\mu_{..2} = \left(\frac{1}{2 \cdot 2 \cdot 3} \right) [9 + 10 + 8 + 35 + 29 + 33 + 34 + 30 + 32 + 12 + 16 + 15] = 21,91$$

Por tanto:

$$\gamma_1 = \mu_{..1} - \mu = 19,25 - 20,58 = -1,33$$

$$\gamma_2 = \mu_{..2} - \mu = 21,91 - 20,58 = 1,33$$

El *vector C* adopta los siguientes valores:

$$C = \{\gamma\} = \begin{bmatrix} -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \\ -1,33 \\ -1,33 \\ -1,33 \\ 1,33 \\ 1,33 \\ 1,33 \end{bmatrix} \quad (7.37)$$

VECTOR $\{\alpha\beta\}$

La variabilidad asociada al vector $\{\alpha\beta\}$ es la correspondiente al efecto de interacción entre los factores *A* y *B*. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\alpha\beta)_{jk}$.

$$(\alpha\beta)_{jk} = \mu_{jk.} - (\mu + \alpha_j + \beta_k) \quad (7.38)$$

donde:

$$\mu_{jk.} = \frac{1}{cn} \sum_m \sum_i Y_{ijkm}$$

Promedios correspondientes a las combinaciones entre los niveles de los factores *A* y *B*:

$$\mu_{11.} = \left(\frac{1}{2 \cdot 3}\right) [22 + 20 + 18 + 9 + 10 + 8] = 14,5$$

$$\mu_{12.} = \left(\frac{1}{2 \cdot 3}\right) [14 + 13 + 14 + 35 + 29 + 33] = 23$$

$$\mu_{21.} = \left(\frac{1}{2 \cdot 3}\right) [16 + 18 + 14 + 34 + 30 + 32] = 24$$

$$\mu_{22.} = \left(\frac{1}{2 \cdot 3}\right) [28 + 29 + 25 + 12 + 16 + 15] = 20,83$$

$$\mu_{2.1} = \left(\frac{1}{2 \cdot 3}\right) [16 + 18 + 14 + 28 + 29 + 25] = 21,66$$

$$\mu_{2.2} = \left(\frac{1}{2 \cdot 3}\right) [34 + 30 + 32 + 12 + 16 + 15] = 23,16$$

Por tanto:

$$(\alpha\gamma)_{11} = \mu_{1.1} - \mu - \alpha_1 - \gamma_1 = 16,83 - 20,58 - (-1,83) - (-1,33) = -0,58$$

$$(\alpha\gamma)_{12} = \mu_{1.2} - \mu - \alpha_1 - \gamma_2 = 20,66 - 20,58 - (-1,83) - 1,33 = 0,58$$

$$(\alpha\gamma)_{21} = \mu_{2.1} - \mu - \alpha_2 - \gamma_1 = 21,66 - 20,58 - 1,83 - (-1,33) = 0,58$$

$$(\alpha\gamma)_{22} = \mu_{2.2} - \mu - \alpha_2 - \gamma_2 = 23,16 - 20,58 - 1,83 - 1,33 = -0,58$$

El vector $\{\alpha\gamma\}$ adopta los siguientes valores:

$$\{\alpha\gamma\} = \begin{bmatrix} -0,58 \\ -0,58 \\ -0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ -0,58 \\ -0,58 \\ -0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ -0,58 \\ -0,58 \\ -0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ 0,58 \\ -0,58 \\ -0,58 \\ -0,58 \end{bmatrix} \quad (7.41)$$

VECTOR $\{\beta\gamma\}$

La variabilidad asociada al vector $\{\beta\gamma\}$ es la correspondiente al efecto de interacción entre los factores B y C . Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\beta\gamma)_{km}$.

$$(\beta\gamma)_{km} = \mu_{.km} - (\mu + \beta_k + \gamma_m)$$

donde:

$$\mu_{.km} = \frac{1}{an} \sum_j \sum_i Y_{ijk} \quad (7.42)$$

Calculando los promedios $\mu_{.km}$ obtenemos:

$$\mu_{.11} = \left(\frac{1}{2 \cdot 3} \right) [22 + 20 + 18 + 16 + 18 + 14] = 18$$

$$\mu_{.12} = \left(\frac{1}{2 \cdot 3} \right) [9 + 10 + 8 + 34 + 30 + 32] = 20,5$$

$$\mu_{.21} = \left(\frac{1}{2 \cdot 3} \right) [14 + 13 + 14 + 28 + 29 + 25] = 20,5$$

$$\mu_{.22} = \left(\frac{1}{2 \cdot 3} \right) [35 + 29 + 33 + 12 + 16 + 15] = 23,33$$

Por tanto:

$$(\beta\gamma)_{11} = \mu_{.11} - \mu - \beta_1 - \gamma_1 = 18 - 20,58 - (-1,33) - (-1,33) = 0,08$$

$$(\beta\gamma)_{12} = \mu_{.12} - \mu - \beta_1 - \gamma_2 = 20,5 - 20,58 - (-1,33) - 1,33 = -0,08$$

$$(\beta\gamma)_{21} = \mu_{.21} - \mu - \beta_2 - \gamma_1 = 20,5 - 20,58 - 1,33 - (-1,33) = -0,08$$

$$(\beta\gamma)_{22} = \mu_{.22} - \mu - \beta_2 - \gamma_2 = 23,33 - 20,58 - 1,33 - 1,33 = 0,08$$

El vector $\{\beta\gamma\}$ adopta los siguientes valores:

$$\{\beta\gamma\} = \begin{bmatrix} 0,08 \\ 0,08 \\ 0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ 0,08 \\ 0,08 \\ 0,08 \\ 0,08 \\ 0,08 \\ 0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ -0,08 \\ 0,08 \\ 0,08 \\ 0,08 \end{bmatrix} \quad (7.43)$$

VECTOR $\{\alpha\beta\gamma\}$

La variabilidad asociada al vector $\{\alpha\beta\gamma\}$ es la correspondiente al efecto de interacción entre los factores A, B y C. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre estos tres factores, $(\alpha\beta\gamma)_{jkm}$:

$$(\alpha\beta\gamma)_{jkm} = \mu_{jkm} - (\mu + \alpha_j + \beta_k + \gamma_m + (\alpha\beta)_{jk} + (\alpha\gamma)_{jm} + (\beta\gamma)_{km}) \quad (7.44)$$

donde:

$$\mu_{jkm} = \frac{1}{n} \sum_i Y_{ijkm}$$

Calculando los promedios correspondientes a las combinaciones entre los niveles de los factores A, B y C, es decir, los μ_{jkm} , obtenemos:

$$\mu_{111} = \left(\frac{1}{3}\right) [22 + 20 + 18] = 20$$

$$\mu_{112} = \left(\frac{1}{3}\right) [9 + 10 + 8] = 9$$

$$\mu_{121} = \left(\frac{1}{3}\right) [14 + 13 + 14] = 13,66$$

$$\mu_{122} = \left(\frac{1}{3}\right) [35 + 29 + 33] = 32,33$$

$$\mu_{211} = \left(\frac{1}{3}\right) [16 + 18 + 14] = 16$$

$$\mu_{212} = \left(\frac{1}{3}\right) [34 + 30 + 32] = 32$$

$$\mu_{221} = \left(\frac{1}{3}\right) [28 + 29 + 25] = 27,33$$

$$\mu_{222} = \left(\frac{1}{3}\right) [12 + 16 + 15] = 14,33$$

Por tanto:

$$(\alpha\beta\gamma)_{111} = \mu_{111} - \mu - \alpha_1 - \beta_1 - \gamma_1 - (\alpha\beta)_{11} - (\alpha\gamma)_{11} - (\beta\gamma)_{11}$$

$$(\alpha\beta\gamma)_{111} = 20 - 20,58 - (-1,83) - (-1,33) - (-1,33) - (-2,91) - (-0,58) - 0,08 = 7,33$$

$$(\alpha\beta\gamma)_{112} = \mu_{112} - \mu - \alpha_1 - \beta_1 - \gamma_2 - (\alpha\beta)_{11} - (\alpha\gamma)_{12} - (\beta\gamma)_{12}$$

$$(\alpha\beta\gamma)_{112} = 9 - 20,58 - (-1,83) - (-1,33) - 1,33 - (-2,91) - 0,58 - (-0,08) = -7,33$$

$$(\alpha\beta\gamma)_{121} = \mu_{121} - \mu - \alpha_1 - \beta_2 - \gamma_1 - (\alpha\beta)_{12} - (\alpha\gamma)_{11} - (\beta\gamma)_{21}$$

$$(\alpha\beta\gamma)_{121} = 13,66 - 20,58 - (-1,83) - 1,33 - (-1,33) - 2,91 - (-0,58) - (-0,08) = -7,33$$

$$(\alpha\beta\gamma)_{122} = \mu_{122} - \mu - \alpha_1 - \beta_2 - \gamma_2 - (\alpha\beta)_{12} - (\alpha\gamma)_{12} - (\beta\gamma)_{22}$$

$$(\alpha\beta\gamma)_{122} = 32,33 - 20,58 - (-1,83) - 1,33 - 1,33 - 2,91 - 0,58 - 0,08 = 7,33$$

$$(\alpha\beta\gamma)_{211} = \mu_{211} - \mu - \alpha_2 - \beta_1 - \gamma_1 - (\alpha\beta)_{21} - (\alpha\gamma)_{21} - (\beta\gamma)_{11}$$

$$(\alpha\beta\gamma)_{211} = 16 - 20,58 - 1,83 - (-1,33) - (-1,33) - 2,91 - 0,58 - 0,08 = -7,33$$

$$(\alpha\beta\gamma)_{212} = \mu_{212} - \mu - \alpha_2 - \beta_1 - \gamma_2 - (\alpha\beta)_{21} - (\alpha\gamma)_{22} - (\beta\gamma)_{12}$$

$$(\alpha\beta\gamma)_{212} = 32 - 20,58 - 1,83 - (-1,33) - 1,33 - 2,91 - (-0,58) - (-0,08) = 7,33$$

$$(\alpha\beta\gamma)_{221} = \mu_{221} - \mu - \alpha_2 - \beta_2 - \gamma_1 - (\alpha\beta)_{22} - (\alpha\gamma)_{21} - (\beta\gamma)_{21}$$

$$(\alpha\beta\gamma)_{221} = 27,33 - 20,58 - 1,83 - 1,33 - (-1,33) - (-2,91) - 0,58 - (-0,08) = 7,33$$

$$(\alpha\beta\gamma)_{222} = \mu_{222} - \mu - \alpha_2 - \beta_2 - \gamma_2 - (\alpha\beta)_{22} - (\alpha\gamma)_{22} - (\beta\gamma)_{22}$$

$$(\alpha\beta\gamma)_{222} = 14,33 - 20,58 - 1,83 - 1,33 - 1,33 - (-2,91) - 0,58 - (-0,08) = -7,33$$

El vector $\{\alpha\beta\gamma\}$ adopta los siguientes valores:

$$\{\alpha\beta\gamma\} = \begin{bmatrix} 7,33 \\ 7,33 \\ 7,33 \\ -7,33 \\ -7,33 \\ -7,33 \\ -7,33 \\ -7,33 \\ 7,33 \\ 7,33 \\ 7,33 \\ -7,33 \\ -7,33 \\ -7,33 \\ 7,33 \\ 7,33 \\ 7,33 \\ 7,33 \\ 7,33 \\ 7,33 \\ -7,33 \\ -7,33 \\ -7,33 \end{bmatrix} \quad (7.45)$$

VECTOR E, CÁLCULO DE LOS RESIDUALES O ERRORES

Partiendo de la Fórmula (7.21):

$$\varepsilon_{ijkm} = Y_{ijkm} - [\mu + \alpha_j + \beta_k + \gamma_m + (\alpha\beta)_{jk} + (\alpha\gamma)_{jm} + (\beta\gamma)_{km} + (\alpha\beta\gamma)_{jkm}]$$

$$\varepsilon_{i111} = Y_{i111} - (\mu + \alpha_1 + \beta_1 + \gamma_1 + (\alpha\beta)_{11} + (\alpha\gamma)_{11} + (\beta\gamma)_{11} + (\alpha\beta\gamma)_{111})$$

$$\begin{aligned} \varepsilon_{i111} &= Y_{i111} - (20,58 + (-1,83) + (-1,33) + (-1,33) + (-2,91) + (-0,58) + \\ &\quad + 0,08 + 7,33) = Y_{i111} - 20 \end{aligned}$$

Sustituyendo las puntuaciones Y_{i111} por sus correspondientes valores:

$$\varepsilon_{1111} = 22 - 20 = 2$$

$$\varepsilon_{2111} = 20 - 20 = 0$$

$$\varepsilon_{3111} = 18 - 20 = -2$$

$$\varepsilon_{i112} = Y_{i112} - (\mu + \alpha_1 + \beta_1 + \gamma_2 + (\alpha\beta)_{11} + (\alpha\gamma)_{12} + (\beta\gamma)_{12} + (\alpha\beta\gamma)_{112})$$

$$\varepsilon_{i112} = Y_{i112} - (20,58 + (-1,83) + (-1,33) + 1,33 + (-2,91) + 0,58 + (-0,08) + (-7,33)) = Y_{i112} - 9$$

Sustituyendo las puntuaciones Y_{i112} por sus correspondientes valores:

$$\varepsilon_{1112} = 9 - 9 = 0$$

$$\varepsilon_{2112} = 10 - 9 = 1$$

$$\varepsilon_{3112} = 8 - 9 = -1$$

$$\varepsilon_{i121} = Y_{i121} - (\mu + \alpha_1 + \beta_2 + \gamma_1 + (\alpha\beta)_{12} + (\alpha\gamma)_{11} + (\beta\gamma)_{21} + (\alpha\beta\gamma)_{121})$$

$$\varepsilon_{i121} = Y_{i121} - (20,58 + (-1,83) + 1,33 + (-1,33) + 2,91 + (-0,58) + (-0,08) + (-7,33)) = Y_{i121} - 13,66$$

Sustituyendo las puntuaciones Y_{i121} por sus correspondientes valores:

$$\varepsilon_{1121} = 14 - 13,66 = 0,33$$

$$\varepsilon_{2121} = 13 - 13,66 = -0,66$$

$$\varepsilon_{3121} = 14 - 13,66 = 0,33$$

$$\varepsilon_{i122} = Y_{i122} - (\mu + \alpha_1 + \beta_2 + \gamma_2 + (\alpha\beta)_{12} + (\alpha\gamma)_{12} + (\beta\gamma)_{22} + (\alpha\beta\gamma)_{122})$$

$$\varepsilon_{i122} = Y_{i122} - (20,58 + (-1,83) + 1,33 + 1,33 + 2,91 + 0,58 + 0,08 + 7,33) = Y_{i122} - 32,33$$

Sustituyendo las puntuaciones Y_{i122} por sus correspondientes valores:

$$\varepsilon_{1122} = 35 - 32,33 = 2,66$$

$$\varepsilon_{2122} = 29 - 32,33 = -3,33$$

$$\varepsilon_{3122} = 33 - 32,33 = 0,66$$

$$\varepsilon_{i211} = Y_{i211} - (\mu + \alpha_2 + \beta_1 + \gamma_1 + (\alpha\beta)_{21} + (\alpha\gamma)_{21} + (\beta\gamma)_{11} + (\alpha\beta\gamma)_{211})$$

$$\varepsilon_{i211} = Y_{i211} - (20,58 + 1,83 + (-1,33) + (-1,33) + 2,91 + 0,58 + 0,08 + (-7,33)) = Y_{i211} - 16$$

Sustituyendo las puntuaciones Y_{i211} por sus correspondientes valores:

$$\varepsilon_{1211} = 16 - 16 = 0$$

$$\varepsilon_{2211} = 18 - 16 = 2$$

$$\varepsilon_{3211} = 14 - 16 = -2$$

$$\varepsilon_{i212} = Y_{i212} - (\mu + \alpha_2 + \beta_1 + \gamma_2 + (\alpha\beta)_{21} + (\alpha\gamma)_{22} + (\beta\gamma)_{12} + (\alpha\beta\gamma)_{212})$$

$$\varepsilon_{i212} = Y_{i212} - (20,58 + 1,83 + (-1,33) + 1,33 + 2,91 + (-0,58) + (-0,08) + 7,33) = Y_{i212} - 32$$

Sustituyendo las puntuaciones Y_{i212} por sus correspondientes valores:

$$\varepsilon_{1212} = 34 - 32 = 2$$

$$\varepsilon_{2212} = 30 - 32 = -2$$

$$\varepsilon_{3212} = 32 - 32 = 0$$

$$\varepsilon_{i221} = Y_{i221} - (\mu + \alpha_2 + \beta_2 + \gamma_1 + (\alpha\beta)_{22} + (\alpha\gamma)_{21} + (\beta\gamma)_{21} + (\alpha\beta\gamma)_{221})$$

$$\varepsilon_{i221} = Y_{i221} - (20,58 + 1,83 + 1,33 + (-1,33) + (-2,91) + 0,58 + (-0,08) + 7,33) = Y_{i221} - 27,33$$

Sustituyendo las puntuaciones Y_{i221} por sus correspondientes valores:

$$\varepsilon_{1221} = 28 - 27,33 = 0,66$$

$$\varepsilon_{2221} = 29 - 27,33 = 1,66$$

$$\varepsilon_{3221} = 25 - 27,33 = -2,33$$

$$\varepsilon_{i222} = Y_{i222} - (\mu + \alpha_2 + \beta_2 + \gamma_2 + (\alpha\beta)_{22} + (\alpha\gamma)_{22} + (\beta\gamma)_{22} + (\alpha\beta\gamma)_{222})$$

$$\varepsilon_{i222} = Y_{i222} - (20,58 + 1,83 + 1,33 + 1,33 + (-2,91) + (-0,58) + 0,08 + (-7,33)) = Y_{i222} - 14,33$$

Sustituyendo las puntuaciones Y_{i222} por sus correspondientes valores:

$$\varepsilon_{1222} = 12 - 14,33 = -2,33$$

$$\varepsilon_{2222} = 16 - 14,33 = 1,66$$

$$\varepsilon_{3222} = 15 - 14,33 = 0,66$$

Por tanto, el vector E adopta los siguientes valores:

$$E = \{\varepsilon\} = \begin{bmatrix} 2 \\ 0 \\ -2 \\ 0 \\ 1 \\ -1 \\ 0,33 \\ -0,66 \\ 0,33 \\ 2,66 \\ -3,33 \\ 0,66 \\ 0 \\ 2 \\ -2 \\ 2 \\ -2 \\ 0 \\ 0,66 \\ 1,66 \\ -2,33 \\ -2,33 \\ 1,66 \\ 0,66 \end{bmatrix} \quad (7.46)$$

Llegados a este punto, podemos calcular, la SCA, la SCB, la SCC, la SCAB, la SCAC, la SCBC, la SCABC y la SCR aplicando las Fórmulas (7.47), (7.48), (7.49), (7.50), (7.51), (7.52), (7.53) y (7.54) o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos.

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = bcn \sum_j \alpha_j^2 = 2 \cdot 2 \cdot 3[(-1,83)^2 + (1,83)^2] = 80,66$$

donde:

$$\begin{aligned}\alpha_j &= \mu_{j..} - \mu \\ \mu &= \frac{1}{a \cdot b \cdot c \cdot n} \sum_j \sum_k \sum_m \sum_i Y_{ijkm} \\ \mu_{j..} &= \frac{1}{b \cdot c \cdot n} \sum_k \sum_m \sum_i Y_{ijkm}\end{aligned}\tag{7.47}$$

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR B (SCB):

$$SCB = acn \sum_k \beta_k^2 = 2 \cdot 2 \cdot 3[(-1,33)^2 + (1,33)^2] = 42,66$$

donde:

$$\begin{aligned}\beta_k &= \mu_{.k.} - \mu \\ \mu_{.k.} &= \frac{1}{a \cdot c \cdot n} \sum_j \sum_m \sum_i Y_{ijkm}\end{aligned}\tag{7.48}$$

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR C (SCC):

$$SCC = abn \sum_m \gamma_m^2 = 2 \cdot 2 \cdot 3[(-1,33)^2 + (1,33)^2] = 42,66$$

donde:

$$\begin{aligned}\gamma_m &= \mu_{..m} - \mu \\ \mu_{..m} &= \frac{1}{a \cdot b \cdot n} \sum_j \sum_k \sum_i Y_{ijkm}\end{aligned}\tag{7.49}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y B (SCAB):

$$SCAB = cn \sum_j \sum_k (\alpha\beta)_{jk}^2 = 2 \cdot 3[(-2,91)^2 + (2,91)^2 + (2,91)^2 + (-2,91)^2] = 204,17$$

donde:

$$\begin{aligned}(\alpha\beta)_{jk} &= \mu_{jk.} - (\mu + \alpha_j + \beta_k) \\ \mu_{jk.} &= \frac{1}{cn} \sum_m \sum_i Y_{ijkm}\end{aligned}\tag{7.50}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y C (SCAC):

$$SCAC = bn \sum_j \sum_m (\alpha\gamma)_{jm}^2 = 2 \cdot 3[(-0,58)^2 + (0,58)^2 + (-0,58)^2 + (-0,58)^2] = 8,17$$

donde:

$$(\alpha\gamma)_{jm} = \mu_{j.m} - (\mu + \alpha_j + \gamma_m) \quad (7.51)$$

$$\mu_{j.m} = \frac{1}{bn} \sum_k \sum_i Y_{ijkm}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES B y C (SCBC):

$$SCBC = an \sum_k \sum_m (\beta\gamma)_{km}^2 = 2 \cdot 3[(0,08)^2 + (-0,08)^2 + (-0,08)^2 + (0,08)^2] = 0,17$$

donde:

$$(\beta\gamma)_{km} = \mu_{.km} - (\mu + \beta_k + \gamma_m) \quad (7.52)$$

$$\mu_{.km} = \frac{1}{an} \sum_j \sum_i Y_{ijkm}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A, B y C (SCABC):

$$SCABC = n \sum_j \sum_k \sum_m (\alpha\beta\gamma)_{jkm}^2 = 3[(7,33)^2 + (-7,33)^2 + (-7,33)^2 + (7,33)^2 + (-7,33)^2 + (7,33)^2 + (7,33)^2 + (-7,33)^2] = 1.290,66$$

donde:

$$(\alpha\beta\gamma)_{jkm} = \mu_{jkm} - (\mu + \alpha_j + \beta_k + \gamma_m + (\alpha\beta)_{jk} + (\alpha\gamma)_{jm} + (\beta\gamma)_{km}) \quad (7.53)$$

$$\mu_{jkm} = \frac{1}{n} \sum_i Y_{ijkm}$$

SUMA CUADRÁTICA CORRESPONDIENTE A LOS RESIDUALES O ERRORES (SCR):

$$SCR = \sum_j \sum_k \sum_m \sum_i \varepsilon_{ijkm}^2 = (2)^2 + (0)^2 + (-2)^2 + \dots + (-2,33)^2 + (1,66)^2 + (0,66)^2 = 62,67$$

donde:

$$\varepsilon_{ijkm} = Y_{ijkm} - [\mu + \alpha_j + \beta_k + \gamma_m + (\alpha\beta)_{jk} + (\alpha\gamma)_{jm} + (\beta\gamma)_{km} + (\alpha\beta\gamma)_{jkm}] \quad (7.54)$$

Multiplicando cada vector por su traspuesto obtenemos los mismos resultados.

$$SCA = \{\alpha\}^T \{\alpha\} = 80,66 \quad (7.55)$$

$$SCAC = \{\alpha\gamma\}^T \{\alpha\gamma\} = 8,17 \quad (7.59)$$

$$SCB = \{\beta\}^T \{\beta\} = 42,66 \quad (7.56)$$

$$SCBC = \{\beta\gamma\}^T \{\beta\gamma\} = 0,17 \quad (7.60)$$

$$SCC = \{\gamma\}^T \{\gamma\} = 42,66 \quad (7.57)$$

$$SCABC = \{\alpha\beta\gamma\}^T \{\alpha\beta\gamma\} = 1.290,66 \quad (7.61)$$

$$SCAB = \{\alpha\beta\}^T \{\alpha\beta\} = 204,17 \quad (7.58)$$

$$SCR = \{\varepsilon\}^T \{\varepsilon\} = 62,67 \quad (7.62)$$

Debemos tener en cuenta que:

$$SCT = SCA + SCB + SCC + SCAB + SCAC + SCBC + SCABC + SCR = 1.731,83$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza (Tabla 7.9).

TABLA 7.9 Análisis factorial de la varianza para el diseño factorial $A \times B \times C$: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A	$SCA = 80,66$	$a - 1 = 1$	$MCA = 80,66$	$F_A = 20,59$
Factor B	$SCB = 42,66$	$b - 1 = 1$	$MCB = 42,66$	$F_B = 10,89$
Factor C	$SCC = 42,66$	$c - 1 = 1$	$MCC = 42,66$	$F_C = 10,89$
Interacción $A \times B$	$SCAB = 204,17$	$(a - 1)(b - 1) = 1$	$MCAB = 204,17$	$F_{AB} = 52,21$
Interacción $A \times C$	$SCAC = 8,17$	$(a - 1)(c - 1) = 1$	$MCAC = 8,17$	$F_{AC} = 2,08$
Interacción $B \times C$	$SCBC = 0,17$	$(b - 1)(c - 1) = 1$	$MCBC = 0,17$	$F_{BC} = 0,04$
Interacción $A \times B \times C$	$SCABC = 1.290,66$	$(a - 1)(b - 1)(c - 1) = 1$	$MCABC = 1.290,66$	$F_{ABC} = 330,09$
Intragrupo, residual o del error	$SCR = 62,67$	$abc(n - 1) = 16$	$MCR = 3,91$	
Total	$SCT = 1.731,83$	$abcn - 1 = 23$		

Tras la obtención de las F observadas, debemos determinar si la variabilidad explicada por los factores y por las interacciones entre tales factores es o no significativa. Para ello, recurrimos a las tablas de los valores críticos de la distribución F . Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los valores críticos que se muestran en la Tabla 7.10.

Como puede observarse en la Tabla 7.10, tanto la *terapia basada en fármacos* (factor A) como la *terapia individual* (factor B) y la *terapia familiar* (factor C) ejercen una influencia estadísticamente significativa sobre el estado físico y psicológico de los pacientes aquejados de SIDA. No obstante, también se observa un efecto de interacción estadísticamente sig-

nificativo entre la *terapia basada en fármacos* y la *terapia individual* (factores *A* y *B*) así como entre los *tres tipos de terapia* (factores *A*, *B* y *C*). Dichas interacciones nos llevan a pensar que, por una parte, las interacciones de primer orden ($A \times B$, $A \times C$ y $B \times C$) no son consistentes y que, además, los efectos principales no se pueden interpretar de manera unívoca.

TABLA 7.10 Comparaciones entre las *F* observadas y las *F* teóricas con un nivel de confianza del 95 %

Fuentes de variación	<i>F</i> crítica _(0,95; gl1/gl2)	<i>F</i> observada	Diferencia
Factor <i>A</i>	$F_{0,95; 1/16} = 4,49$	$F_A = 20,59$	$20,59 > 4,49$
Factor <i>B</i>	$F_{0,95; 1/16} = 4,49$	$F_B = 10,89$	$10,89 > 4,49$
Factor <i>C</i>	$F_{0,95; 1/16} = 4,49$	$F_C = 10,89$	$10,89 > 4,49$
Interacción $A \times B$	$F_{0,95; 1/16} = 4,49$	$F_{AB} = 52,21$	$52,21 > 4,49$
Interacción $A \times C$	$F_{0,95; 1/16} = 4,49$	$F_{AC} = 2,08$	$2,08 < 4,49$
Interacción $B \times C$	$F_{0,95; 1/16} = 4,49$	$F_{BC} = 0,04$	$0,04 < 4,49$
Interacción $A \times B \times C$	$F_{0,95; 1/16} = 4,49$	$F_{ABC} = 330,09$	$330,09 > 4,49$

Por otra parte, se observa que las interacciones $A \times C$ y $B \times C$ no son estadísticamente significativas. Supongamos que en lugar de observar este hecho *a posteriori*, se establece *a priori* que tales interacciones no son aceptables desde un punto de vista teórico. Es decir, postulamos que no existe un fundamento teórico que nos lleve a pensar que pueda producirse algún tipo de interacción entre los factores *A* y *C*, y entre los factores *B* y *C*. En tal caso, la variabilidad correspondiente a dichas interacciones pasaría a formar parte de la variabilidad residual, de forma que:

$$SCR_b = SCAC + SCBC + SCR$$

Así:

$$SCT = SCA + SCB + SCC + SCAB + SCABC + SCR_b$$

Modificando la Tabla 7.9, en función de este nuevo supuesto, obtendríamos la Tabla 7.11.

Estableciendo el mismo nivel de confianza que en el caso anterior, obtenemos los valores críticos que se muestran en la Tabla 7.12.

Comparando la Tabla 7.10 con la Tabla 7.12 se aprecia que, al eliminar del modelo aquellas interacciones que no son aceptables desde el punto de vista teórico, aumenta la potencia probatoria del diseño y, por tanto, disminuye la probabilidad de cometer un error de tipo II. En consecuencia, consideramos que el no incluir las interacciones que carecen de base teórica en el modelo, es una práctica muy recomendable.

De la misma forma, aunque en todos los ejemplos prácticos del presente texto se calculan los efectos asociados a los parámetros que configuran los modelos no aditivos de los distintos diseños, hemos de reiterar que en caso de que las interacciones entre los factores no sean estadísticamente significativas, los modelos aditivos representan de forma más adecuada la realidad.

TABLA 7.11 Análisis factorial de la varianza para el diseño factorial $A \times B \times C$: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A	$SCA = 80,66$	$a - 1 = 1$	$MCA = 80,66$	$F_A = 20,47$
Factor B	$SCB = 42,66$	$b - 1 = 1$	$MCB = 42,66$	$F_B = 10,83$
Factor C	$SCC = 42,66$	$c - 1 = 1$	$MCC = 42,66$	$F_C = 10,83$
Interacción $A \times B$	$SCAB = 204,17$	$(a - 1)(b - 1) = 1$	$MCAB = 204,17$	$F_{AB} = 51,82$
Interacción $A \times B \times C$	$SCABC = 1.290,66$	$(a - 1)(b - 1)(c - 1) = 1$	$MCABC = 1.290,66$	$F_{ABC} = 327,58$
Intragrupo, residual o del error	$SCR_b = 71,01$	$gl_{total} - gl_A - gl_B - gl_C - gl_{AB} - gl_{ABC} = 18$	$MCR = 3,94$	
Total	$SCT = 1.731,83$	$abcn - 1 = 23$		

TABLA 7.12 Comparaciones entre las F observadas y las F teóricas con un nivel de confianza del 95 %

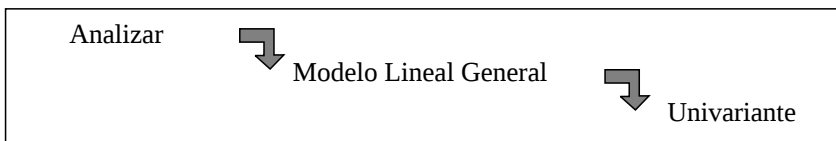
Fuentes de variación	F crítica _(0,95; gl1/gl2)	F observada	Diferencia
Factor A	$F_{0,95; 1/18} = 4,41$	$F_A = 20,47$	$20,47 > 4,41$
Factor B	$F_{0,95; 1/18} = 4,41$	$F_B = 10,83$	$10,83 > 4,41$
Factor C	$F_{0,95; 1/18} = 4,41$	$F_C = 10,83$	$10,83 > 4,41$
Interacción $A \times B$	$F_{0,95; 1/18} = 4,41$	$F_{AB} = 51,82$	$51,82 > 4,41$
Interacción $A \times B \times C$	$F_{0,95; 1/18} = 4,41$	$F_{ABC} = 327,58$	$327,58 > 4,41$

7.3.3.4. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

Análisis factorial de la varianza

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la(s) variable(s) dependiente(s) y los factores, especificando en su lugar correspondiente si tales factores son fijos o aleatorios (en nuestro ejemplo, los tres factores son fijos).



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor así como para la interacción. Asimismo, permite seleccionar de entre un conjunto de opciones aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar* «estadísticos descriptivos», «estimaciones del tamaño del efecto», «potencia observada» y «pruebas de homogeneidad»).



- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

UNIANOVA

```
estado BY farmaco individu familiar
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(farmaco)
/EMMEANS = TABLES(individu)
/EMMEANS = TABLES(familiar)
/PRINT = DESCRIPTIVE ETASQ OPOWER HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = farmaco individu familiar farmaco*individu farmaco*familiar
individu*familiar farmaco*individu*familiar .
```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Terapia basada en fármacos	1,00	No recibe terapia	12
	2,00	Recibe terapia	12
Terapia individual	1,00	No	12
	2,00	Sí	12
Terapia familiar	1,00	No	12
	2,00	Sí	12

Estadísticos descriptivos

Variable dependiente: Estado físico y psicológico

T. fármacos	T. individual	T. familiar	Media	Desv. típ.	N
No recibe terapia	No	No	20,0000	2,0000	3
		Sí	9,0000	1,0000	3
		Total	14,5000	6,1887	6
	Sí	No	13,6667	0,5774	3
		Sí	32,3333	3,0551	3
		Total	23,0000	10,4115	6
	Total	No	16,8333	3,7103	6
		Sí	20,6667	12,9409	6
		Total	18,7500	9,2944	12

Variable dependiente: Estado físico y psicológico (continuación)

T. fármacos	T. individual	T. familiar	Media	Desv. típ.	N
Recibe terapia	No	No	16,0000	2,0000	3
		Sí	32,0000	2,0000	3
		Total	24,0000	8,9443	6
	Sí	No	27,3333	2,0817	3
		Sí	14,3333	2,0817	3
		Total	20,8333	7,3598	6
	Total	No	21,6667	6,4704	6
		Sí	23,1667	9,8472	6
		Total	22,4167	7,9825	12
Total	No	No	18,0000	2,8284	6
		Sí	20,5000	12,6768	6
		Total	19,2500	8,8536	12
	Sí	No	20,5000	7,6092	6
		Sí	23,3333	10,1325	6
		Total	21,9167	8,6703	12
	Total	No	19,2500	5,6266	12
		Sí	21,9167	11,0409	12
		Total	20,5833	8,6774	24

Contraste de Levene sobre la igualdad de las varianzas error^a

Variable dependiente: Estado físico y psicológico

F	gl1	gl2	Sig.
0,942	7	16	0,502

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a lo largo de todos los grupos.

^a Diseño: Intercept + FARMACO + INDIVIDU + FAMILIAR + FARMACO * INDIVIDU + FARMACO * FAMILIAR + INDIVIDU * FAMILIAR + FARMACO * INDIVIDU * FAMILIAR

Pruebas de los efectos intersujetos**Variable dependiente: Estado físico y psicológico**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	1.669,167 ^b	7	238,452	60,881	0,000	0,964	426,170	1,000
Intercept	10.168,167	1	10.168,167	2.596	0,000	0,994	2.596,1	1,000
FARMACO	80,667	1	80,667	20,596	0,000	0,563	20,596	0,989
INDIVIDU	42,667	1	42,667	10,894	0,005	0,405	10,894	0,872
FAMILIAR	42,667	1	42,667	10,894	0,005	0,405	10,894	0,872
FARMACO * INDIVIDU	204,167	1	204,167	52,128	0,000	0,765	52,128	1,000
FARMACO * FAMILIAR	8,167	1	8,167	2,085	0,168	0,115	2,085	0,274
INDIVIDU * FAMILIAR	0,167	1	0,167	0,043	0,839	0,003	0,043	0,054
FARMACO * INDIVIDU * FAMILIAR	1.290,667	1	1.290,667	329,5	0,000	0,954	329,532	1,000
Error	62,667	16	3,917					
Total	11.900,000	24						
Total corregido	1.731,833	23						

^a Calculado con $\alpha = 0,05$.^b R^2 cuadrado = 0,964 (R^2 cuadrado corregido = 0,948).**Medias marginales estimadas****1. Terapia basada en fármacos****Variable dependiente: Estado físico y psicológico**

Terapia basada en fármacos	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
No recibe terapia	18,750	0,571	17,539	19,961
Recibe terapia	22,417	0,571	21,206	23,628

2. Terapia individual**Variable dependiente: Estado físico y psicológico**

Terapia individual	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
No	19,250	0,571	18,039	20,461
Sí	21,917	0,571	20,706	23,128

3. Terapia familiar

Variable dependiente: Estado físico y psicológico

Terapia familiar	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
No	19,250	0,571	18,039	20,461
Sí	21,917	0,571	20,706	23,128

7.3.4. El estudio de los efectos de interacción:
Análisis de los efectos simples

Teniendo presente la polémica que existe al respecto, y que ya ha sido abordada en el Epígrafe 7.2, cabe afirmar que la estrategia más utilizada entre los metodólogos para interpretar una interacción significativa consiste en el **análisis de los efectos simples**. Riba (1990) define los *efectos simples* como los efectos atribuidos a un componente del modelo lineal, después de fijar los niveles de otro componente de dicho modelo. Por tanto, tales efectos no hacen referencia a todas las observaciones incluidas en el diseño, sino a un subconjunto de las mismas. El análisis de los efectos simples consiste en contrastar, de forma independiente, los efectos de un factor, en cada uno de los niveles del resto de los factores incluidos en el diseño. Así, por ejemplo, en el caso de un diseño factorial $A \times B$, el contraste de los efectos simples del factor A supone considerar su efecto en cada uno de los niveles del factor B , existiendo tantos efectos simples del factor A como niveles tenga el factor B . Lo mismo cabe decir con respecto al factor B . Cada uno de estos contrastes se ejecuta de manera independiente a través de diseños simples.

En caso de que las varianzas de todas las condiciones experimentales sean homogéneas, Maxwell y Delaney (1990) recomiendan utilizar como denominador de la razón F , correspondiente a cada uno de los contrastes, la media cuadrática del error del análisis de la varianza realizado previamente. Si bien en tal circunstancia, el hecho de tomar la varianza poblacional como denominador de la razón F permite maximizar la potencia estadística de la prueba; cuando las varianzas no son homogéneas resulta más adecuado tomar, como término de error, el correspondiente a los grupos implicados en el contraste. Por otra parte, como señala Pascual (1995a), la realización de contrastes de efectos simples que implican un factor con más de dos niveles requiere llevar a cabo comparaciones individuales o contrastes específicos entre las medias de las «celdillas», con el objetivo de dilucidar entre qué niveles del factor existen diferencias estadísticamente significativas.

En este apartado, abordaremos el cálculo y la interpretación de los efectos simples partiendo de una matriz de datos hipotética, correspondiente a un experimento realizado mediante un diseño factorial $A \times B$. Arnau (1986), Pascual (1995a) y Pascual, García y Frías (1995), entre otros, proporcionan varios ejemplos en los que nos muestran diferentes estrategias para llevar a cabo el análisis de los efectos simples en distintos modelos de diseño. El lector interesado en profundizar en la interpretación de la interacción, a partir de tales efectos, puede consultar los textos citados.

7.3.4.1. Ejemplo práctico

Supongamos que, en el ámbito de la psicología educativa, realizamos una investigación para examinar la influencia que ejercen la *actitud del profesor* (factor A) y la *metodología de trabajo empleada en el aula* (factor B) sobre el *rendimiento presentado por un grupo de niños en la asignatura de Ciencias Naturales*. Para ello, seleccionamos aleatoriamente 12 niños de un centro educativo. La mitad de la muestra, escogida al azar, recibe la enseñanza impartida por un *profesor autoritario* (a_1) y la otra mitad se somete a la tutela de un *profesor liberal* (a_2). A su vez, los sujetos se subdividen, aleatoriamente, en tres grupos, en función de la metodología de trabajo empleada en el aula, a saber: (b_1) *metodología tradicional*, (b_2) *metodología basada en el contacto con la naturaleza* y (b_3) *metodología mixta* o consistente en una combinación entre los dos métodos anteriores. En la Tabla 7.13 pueden observarse los resultados obtenidos en la investigación.

TABLA 7.13 Matriz de datos del experimento

		B (Metodología de trabajo)		
		b_1 (Tradicional)	b_2 (Contacto con la naturaleza)	b_3 (Mixta)
A (Actitud del profesor)	a_1 (Autoritaria)	40 44	13 11	8 16
	a_2 (Liberal)	6 10	13 27	10 30

Las tablas de medias y de efectos estimados⁵, para cada una de las celdillas del diseño, adoptan los siguientes valores:

TABLA 7.14 Medias de las condiciones experimentales del diseño factorial Actitud del profesor × Metodología de trabajo

		B			Medias marginales
		b_1	b_2	b_3	
A	a_1	42	12	12	22
	a_2	8	20	20	16
Medias marginales		25	16	16	19

⁵ Mostramos el cálculo de tres efectos estimados, a fin de recordar al lector cómo se obtienen tales efectos:

$$\alpha_1 = \mu_{1.} - \mu = 22 - 19 = 3$$

$$\beta_1 = \mu_{.1} - \mu = 25 - 19 = 6$$

$$(\alpha\beta)_{11} = \mu_{11} - \mu - \alpha_1 - \beta_1 = 42 - 19 - 3 - 6 = 14$$

TABLA 7.15 Efectos estimados para las celdillas del diseño factorial (Actitud del profesor × Metodología de trabajo)

		<i>B</i>			Efectos marginales
		β_1	β_2	β_3	
<i>A</i>	α_1	14	− 7	− 7	3
	α_2	− 14	7	7	− 3
	Efectos marginales	6	− 3	− 3	

Recordemos que la ecuación estructural del ANOVA para el diseño factorial $A \times B$ responde a la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \tag{7.1}$$

Por su parte, los vectores asociados a los parámetros de dicha ecuación adoptan los siguientes valores:

$$Y = \{Y\} = \begin{bmatrix} 40 \\ 44 \\ 6 \\ 10 \\ 13 \\ 13 \\ 11 \\ 13 \\ 27 \\ 8 \\ 16 \\ 10 \\ 30 \end{bmatrix} \tag{7.63}$$

$$\mu = \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} \tag{7.64}$$

$$A = \{\alpha\} = \mu_{j.} - \mu \Rightarrow \{\alpha\} = \begin{bmatrix} 3 \\ 3 \\ -3 \\ -3 \\ 3 \\ 3 \\ -3 \\ -3 \\ 3 \\ 3 \\ -3 \\ -3 \end{bmatrix} \tag{7.65}$$

$$B = \{\beta\} = \mu_{.k} - \mu \Rightarrow \{\beta\} = \begin{bmatrix} 6 \\ 6 \\ 6 \\ 6 \\ -3 \\ -3 \\ -3 \\ -3 \\ -3 \\ -3 \\ -3 \\ -3 \end{bmatrix} \tag{7.66}$$

$$AB = \{\alpha\beta\} = \mu_{jk} - (\mu + \alpha_j + \beta_k) \Rightarrow \{\alpha\beta\} = \begin{bmatrix} 14 \\ 14 \\ -14 \\ -14 \\ -7 \\ -7 \\ 7 \\ 7 \\ -7 \\ -7 \\ 7 \\ 7 \end{bmatrix} \quad (7.67)$$

$$E = \{\varepsilon\} = y_{ijk} - \mu - \alpha_j - \beta_k - (\alpha\beta)_{jk} \Rightarrow \{\varepsilon\} = \begin{bmatrix} -2 \\ 2 \\ -2 \\ 2 \\ 1 \\ -1 \\ -7 \\ 7 \\ -4 \\ 4 \\ -10 \\ 10 \end{bmatrix} \quad (7.68)$$

Las sumas de cuadrados de las diferentes fuentes de variación son:

$$SCA = \{\alpha\}^T \{\alpha\} = 108 \quad (7.69)$$

$$SCB = \{\beta\}^T \{\beta\} = 216 \quad (7.70)$$

$$SCAB = \{\alpha\beta\}^T \{\alpha\beta\} = 1.176 \quad (7.71)$$

$$SCR = \{\varepsilon\}^T \{\varepsilon\} = 348 \quad (7.72)$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 7.16 Análisis factorial de la varianza para el diseño factorial (Actitud del profesor $\times$ Metodología de trabajo: ejemplo práctico)

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F	p
Factor A	$SCA = 108$	$a - 1 = 1$	$MCA = 108$	$F_A = 1,86$	$p > 0,05$
Factor B	$SCB = 216$	$b - 1 = 2$	$MCB = 108$	$F_B = 1,86$	$p > 0,05$
Interacción A $\times$ B	$SCAB = 1.176$	$(a - 1)(b - 1) = 2$	$MCAB = 588$	$F_{AB} = 10,14$	$p < 0,05$
Intragrupo, residual o del error	$SCR = 348$	$ab(n - 1) = 6$	$MCR = 58$		
Total	$SCT = 1.848$	$abn - 1 = 11$			

Los resultados del ANOVA nos llevan a aceptar la H_0 , tanto con respecto al efecto principal de la *actitud del profesor* (factor A), como de la *metodología de trabajo empleada en el aula* (factor B). Por tanto, considerados independientemente, ninguno de los dos factores ejerce una influencia, estadísticamente significativa, sobre *el rendimiento presentado por los niños en la asignatura de Ciencias Naturales*. No obstante, se observa un efecto de interacción significativo entre los factores A y B .

Procedamos a interpretar este efecto de interacción a partir de la representación gráfica de los promedios obtenidos por los sujetos en las diferentes condiciones experimentales (véase la Figura 7.1).

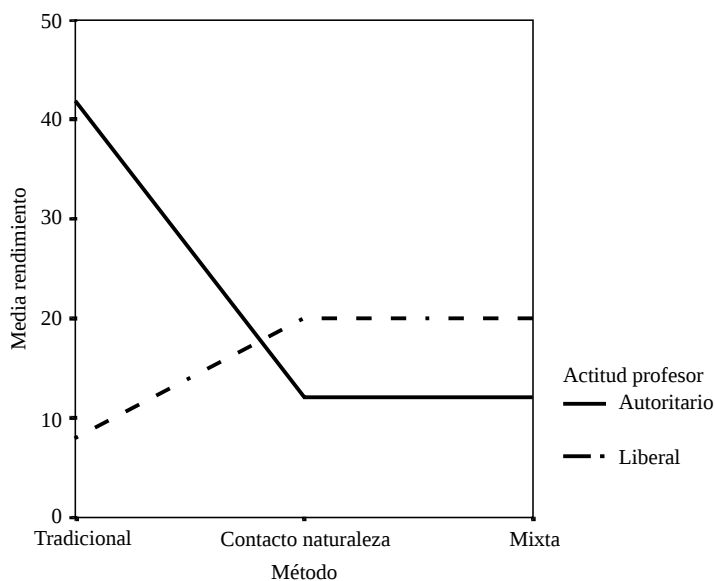


Figura 7.1 Representación gráfica de la interacción $A \times B$.

La representación gráfica pone de manifiesto que, cuando se utiliza una metodología de trabajo tradicional, la actitud autoritaria por parte del profesor lleva a un mejor rendimiento que la actitud liberal. Sin embargo, cuando se emplea un método basado en el contacto con la naturaleza o un método mixto, es el profesor liberal el que obtiene mejor rendimiento por parte de sus alumnos, aunque estas diferencias no parecen ser de suficiente magnitud como para alcanzar la significación estadística. Por otra parte, también cabe explicar el efecto de interacción argumentando que, cuando el profesor presenta una actitud autoritaria, el rendimiento de los niños que utilizan la metodología tradicional es significativamente superior al de los que emplean los otros dos tipos de métodos. Sin embargo, cuando éstos se hallan bajo la tutela del docente liberal, las diferencias existentes en el rendimiento en función de la metodología de trabajo empleada en el aula no se vislumbran de manera tan evidente. Es importante señalar que el conjunto de impresiones derivadas de la representación gráfica debe ser contrastado estadísticamente. En caso contrario, tales impresiones podrían inducir a errores de interpretación.

Dada la existencia de un efecto de interacción significativo entre los factores A y B , procederemos a su interpretación mediante el análisis de los efectos simples. Para ello,

reducimos el diseño factorial 2×3 a 5 diseños simples. Los grados de libertad, de cada uno de los efectos, se calculan *restando una unidad a la cantidad de grupos del factor general* y, los del error, mediante la expresión $ab(n-1)$. Utilizaremos el procedimiento vectorial para obtener las sumas de cuadrados de los diferentes efectos simples.

$$\alpha(b_1) = \mu_{\alpha(b_1)} - \mu - \beta_1 = \begin{bmatrix} 42 \\ 42 \\ 8 \\ 8 \end{bmatrix} - \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} - \begin{bmatrix} 6 \\ 6 \\ 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 17 \\ 17 \\ -17 \\ -17 \end{bmatrix} \quad (7.73)$$

$$SCA(b_1) = \{\alpha(b_1)\}^T \{\alpha(b_1)\} = 1.156 \quad (7.74)$$

$$\alpha(b_2) = \mu_{\alpha(b_2)} - \mu - \beta_2 = \begin{bmatrix} 12 \\ 12 \\ 20 \\ 20 \end{bmatrix} - \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} - \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \\ 4 \\ 4 \end{bmatrix} \quad (7.75)$$

$$SCA(b_2) = \{\alpha(b_2)\}^T \{\alpha(b_2)\} = 64 \quad (7.76)$$

$$\alpha(b_3) = \mu_{\alpha(b_3)} - \mu - \beta_3 = \begin{bmatrix} 12 \\ 12 \\ 20 \\ 20 \end{bmatrix} - \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} - \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ -4 \\ 4 \\ 4 \end{bmatrix} \quad (7.77)$$

$$SCA(b_3) = \{\alpha(b_3)\}^T \{\alpha(b_3)\} = 64 \quad (7.78)$$

$$\beta(a_1) = \mu_{\beta(a_1)} - \mu - \alpha_1 = \begin{bmatrix} 42 \\ 42 \\ 12 \\ 12 \\ 12 \\ 12 \end{bmatrix} - \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 20 \\ 20 \\ -10 \\ -10 \\ -10 \\ -10 \end{bmatrix} \quad (7.79)$$

$$SCB(a_1) = \{\beta(a_1)\}^T \{\beta(a_1)\} = 1.200 \quad (7.80)$$

$$\beta(a_2) = \mu_{\beta(a_2)} - \mu - \alpha_2 = \begin{bmatrix} 8 \\ 8 \\ 20 \\ 20 \\ 20 \\ 20 \end{bmatrix} - \begin{bmatrix} 19 \\ 19 \\ 19 \\ 19 \\ 19 \\ 19 \end{bmatrix} - \begin{bmatrix} -3 \\ -3 \\ -3 \\ -3 \\ -3 \\ -3 \end{bmatrix} = \begin{bmatrix} -8 \\ -8 \\ 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} \quad (7.81)$$

$$SCB(a_2) = \{\beta(a_2)\}^T \{\beta(a_2)\} = 192 \quad (7.82)$$

La tabla resumen del análisis de los efectos simples de los factores *A* y *B* adopta los siguientes valores:

TABLA 7.17 Resumen del análisis de los efectos simples del diseño factorial 2×3 (Actitud del profesor $\times$ Metodología de trabajo)

Efectos simples	Sumas de cuadrados	Grados de libertad del efecto	Grados de libertad del error	MC_{efecto}	MC_{error}	F	p
<i>A</i> en b_1	1.156	1	6	1.156	58	19,93	$p < 0,05$
<i>A</i> en b_2	64	1	6	64	58	1,10	$p > 0,05$
<i>A</i> en b_3	64	1	6	64	58	1,10	$p > 0,05$
<i>B</i> en a_1	1.200	2	6	600	58	10,34	$p < 0,05$
<i>B</i> en a_2	192	2	6	96	58	1,65	$p > 0,05$
Total	2.676						

Partiendo de las sumas cuadráticas de la Tabla 7.17 y de la Tabla 7.16, cabe comprobar que:

$$\Sigma SCA(b) = SCA + SCAB \Rightarrow 1.156 + 64 + 64 = 108 + 1.176$$

$$\Sigma SCB(a) = SCB + SCAB \Rightarrow 1.200 + 192 = 216 + 1.176$$

De aquí se deduce que la variabilidad total, obtenida en el análisis de los efectos simples, incluye dos componentes de la interacción de forma que, si eliminamos uno de ellos, la variabilidad restante corresponde a la suma cuadrática intergrupos (es decir, $SC_{\text{total}} - SC_{\text{error}}$) del diseño original, a saber: $2.676 - 1.176 = 108 + 216 + 1.176 = 1.500$. Esta igualdad permite comprobar que el análisis se ha realizado correctamente.

Los resultados obtenidos, en el análisis de los efectos simples, ponen de manifiesto que dos de tales efectos son estadísticamente significativos: el efecto de *A* en b_1 y el efecto de *B* en a_1 .

La interpretación del primero de ellos no requiere realizar ningún análisis adicional ya que el factor *A* sólo consta de dos niveles y, en consecuencia, las diferencias observadas en el nivel b_1 del factor *B*, entre los promedios de a_1 y de a_2 , son las que explican el hecho de que este efecto haya resultado estadísticamente significativo. Así, cabe concluir que, cuando se utiliza una metodología de trabajo tradicional, la actitud autoritaria por parte del profesor ($\bar{y}_{11} = 42$) lleva a un mejor rendimiento que la actitud liberal ($\bar{y}_{21} = 8$). Sin embargo, el factor *B* consta de tres categorías, de forma que si un efecto simple resulta estadísticamente significativo, en un determinado nivel de *A*, es necesario formular contrastes individuales entre las medias de las distintas celdillas, a fin de saber entre qué niveles concretos de *B* se producen tales diferencias. Si el diseño es equilibrado, dichos contrastes se llevan a cabo analizando la variabilidad explicada por la razón F_{Ψ} . Procedamos al análisis de los contrastes pertinentes al diseño que nos ocupa.

TABLA 7.18 Medias y contrastes específicos para el diseño factorial 2×3 (Actitud del profesor $\times$ Metodología de trabajo)

Combinación	Media	Contraste (1)	Contraste (2)	Contraste (3)
a_1b_1	42	1	1	0
a_1b_2	12	-1	0	1
a_1b_3	12	0	-1	-1
a_2b_1	8	0	0	0
a_2b_2	20	0	0	0
a_2b_3	20	0	0	0

Recordemos que las sumas de cuadrados de los contrastes se calculan mediante la siguiente expresión:

$$SC_{\Psi} = \frac{(C'M_{ab})^2}{C'C} \quad (7.83)$$

donde C' corresponde al vector fila de coeficientes de la matriz de la hipótesis y M_{ab} al vector de medias de las diferentes condiciones experimentales (véase el subapartado «B. Principales procedimientos de comparaciones múltiples: Ejemplos prácticos» dentro del Epígrafe 6.2.2.4 del Capítulo 6).

Aplicando esta fórmula para calcular las sumas de cuadrados correspondientes a los tres contrastes de nuestro ejemplo práctico, obtenemos:

- Primer contraste: $a_1b_1 | a_1b_2$

$$C'_1 M_{ab} = (1 \ 1 \ -1 \ -1 \ 0 \ 0) \begin{pmatrix} 42 \\ 42 \\ 12 \\ 12 \\ 12 \\ 12 \end{pmatrix} = 60 \quad C'_1 C_1 = (1 \ 1 \ -1 \ -1 \ 0 \ 0) \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = 4$$

$$SC_{\Psi_1} = \frac{(C'_1 M_{ab})^2}{C'_1 C_1} = \frac{(60)^2}{4} = 900$$

- Segundo contraste: $a_1b_1 | a_1b_3$

$$C'_2 M_{ab} = (1 \ 1 \ 0 \ 0 \ -1 \ -1) \begin{pmatrix} 42 \\ 42 \\ 12 \\ 12 \\ 12 \\ 12 \end{pmatrix} = 60 \quad C'_2 C_2 = (1 \ 1 \ 0 \ 0 \ -1 \ -1) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ -1 \\ -1 \end{pmatrix} = 4$$

$$SC_{\Psi_2} = \frac{(C'_2 M_{ab})^2}{C'_2 C_2} = \frac{(60)^2}{4} = 900$$

- Tercer contraste: $a_1b_2 | a_1b_3$

$$C'_3M_{ab} = (0 \ 0 \ 1 \ 1 \ -1 \ -1) \begin{pmatrix} 42 \\ 42 \\ 12 \\ 12 \\ 12 \\ 12 \end{pmatrix} = 0 \quad C'_3C_3 = (0 \ 0 \ 1 \ 1 \ -1 \ -1) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} = 4$$

$$SC_{\Psi_3} = \frac{(C'_3M_{ab})^2}{C'_3C_3} = \frac{(0)^2}{4} = 0$$

Tras calcular las sumas cuadráticas, debemos hallar las razones F de los diferentes contrastes. Para ello, disponemos de la siguiente expresión:

$$F_{\Psi} = \frac{\frac{SC_{\Psi}}{gl_{\Psi}}}{\frac{SC_e}{gl_e}} = \frac{MC_{\Psi}}{MC_e} \quad (7.84)$$

En nuestro caso:

$$F_{\Psi_1} = \frac{900}{58} = 15,52 \quad F_{\Psi_2} = \frac{900}{58} = 15,52 \quad F_{\Psi_3} = \frac{0}{58} = 0$$

Aceptando un error de tipo I de 0,05 para la decisión estadística, cabe concluir que existen diferencias estadísticamente significativas entre b_1 y b_2 en el nivel a_1 del factor A ($F_{\Psi_1} = 15,52 > F_{0,95; 2,6} = 5,143$). Así, podemos afirmar que, cuando el docente muestra una actitud autoritaria, se observan diferencias estadísticamente significativas entre el rendimiento de los alumnos que emplean una metodología de trabajo tradicional ($\bar{y}_{11} = 42$) y el de los que utilizan una metodología basada en el contacto con la naturaleza ($\bar{y}_{12} = 12$). Por otra parte, también se aprecian diferencias estadísticamente significativas entre las categorías b_1 y b_3 bajo el mismo nivel del factor A ($F_{\Psi_2} = 15,52 > F_{0,95; 2,6} = 5,143$). Ello nos lleva a concluir que, ante un profesor autoritario, el rendimiento en la asignatura de Ciencias Naturales es significativamente superior en los niños que emplean el método tradicional ($\bar{y}_{11} = 42$) con respecto a los que trabajan mediante el método mixto ($\bar{y}_{13} = 12$). Por último, cabe señalar que no se observan diferencias estadísticamente significativas entre el rendimiento de los niños que utilizan un método basado en el contacto con la naturaleza ($\bar{y}_{12} = 12$) y el de los que emplean un método mixto ($\bar{y}_{13} = 12$) cuando la actitud del profesor es autoritaria ($F_{\Psi_3} = 0 < F_{0,95; 2,6} = 5,143$).

Como hemos visto en el presente ejemplo, la interpretación de la interacción a través del análisis de los efectos simples requiere, en ocasiones, realizar múltiples contrastes entre las medias de las celdillas, lo que incrementa la probabilidad de cometer un error de tipo I en la decisión estadística. Con el objetivo de mantener la tasa de error de tipo I establecida a priori por el investigador, Maxwell y Delaney (1990) proponen aplicar la corrección de Bonferroni teniendo en cuenta el número de contrastes que pueden efectuarse entre las diferentes categorías del factor, cuyos niveles se pretenden comparar. Así, por ejemplo, en un diseño

factorial 2×3 , el análisis de los efectos simples del factor A requiere corregir el α mediante la siguiente razón: $\alpha = 0,05/3$, mientras que para el análisis de los efectos simples de B hemos de utilizar un $\alpha = 0,05/2$. De este modo, si en nuestro ejemplo queremos mantener el error de tipo I especificado a priori por el investigador, debemos asumir errores de tipo I iguales a 0,016 y 0,025 para los factores A y B , respectivamente. El lector puede comprobar que la conclusión sigue siendo la misma tanto para el efecto simple de A en b_1 ($F_{\text{crít.}; 0,016, 1, 6} = 13,74 < F_{\text{empírica}} = 19,93$) como para el efecto simple de B en a_1 ($F_{\text{crít.}; 0,025, 2, 6} = 7,26 < F_{\text{empírica}} = 10,34$).

Para finalizar con el presente subapartado, proporcionamos un esquema que puede resultar útil para tomar decisiones con respecto al análisis de los efectos principales y de los efectos de interacción en un diseño factorial $A \times B$ (véase la Figura 7.2).

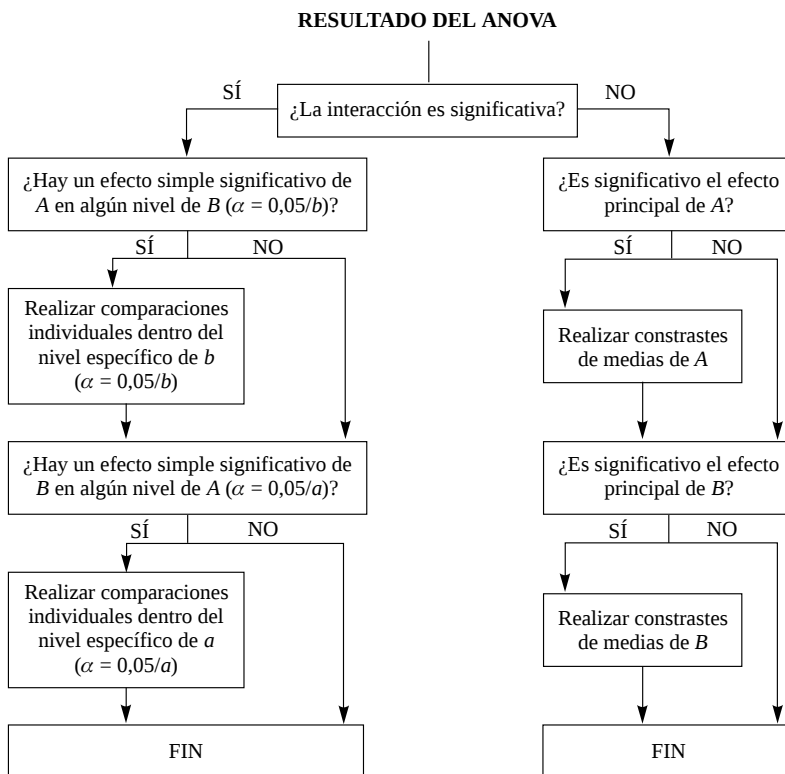


Figura 7.2 Toma de decisiones con respecto al análisis de los efectos principales y de los efectos de interacción en un diseño factorial $A \times B$ (adaptada de Maxwell y Delaney, 1990 en Pascual, García y Frías, 1995, pág. 428).

7.3.5. Comparaciones múltiples entre medias para diseños factoriales con efecto de interacción

Al igual que cuando aplicamos diseños unifactoriales, en el caso de rechazar la hipótesis nula con respecto al efecto de interacción en el diseño factorial, debemos llevar a cabo con-

trastes específicos para determinar entre qué medias existen diferencias estadísticamente significativas tratando de controlar la tasa de error por experimento. La única diferencia con relación al diseño unifactorial es que, cuando trabajamos con un modelo de diseño que incluye un efecto de interacción, la cantidad posible de grupos a comparar es igual a la cantidad de combinaciones posibles entre los niveles de los factores principales, a saber, al número de «celdillas» de las que consta el diseño.

A modo ilustrativo, calcularemos el *rango crítico entre pares de medias* aplicando tres procedimientos de comparaciones múltiples de amplio uso en el ámbito de las ciencias del comportamiento, y determinaremos entre qué medias se producen diferencias estadísticamente significativas tomando como objeto de análisis los datos correspondientes a la interacción que ha resultado significativa en el diseño factorial Actitud del profesor $\times$ Metodología de trabajo (véase la Figura 7.1 en el subapartado «El estudio de los efectos de interacción: Análisis de los efectos simples»). Normalmente, las conclusiones derivadas de las comparaciones múltiples coinciden con las obtenidas a través del análisis de los efectos simples. No obstante, en ocasiones se producen divergencias que se deben al tipo de prueba que se utiliza para efectuar los contrastes.

7.3.5.1. Corrección de Bonferroni

Supongamos que antes de realizar la investigación, el investigador formula tres hipótesis que considera relevantes para examinar el posible efecto de interacción presente en el diseño. La primera de ellas postula que, cuando se utilice una metodología de trabajo tradicional, se producirán diferencias estadísticamente significativas en el rendimiento de los niños en Ciencias Naturales en función de la actitud autoritaria o liberal que muestre el profesor en el aula ($H_0 \equiv \mu_{11} = \mu_{21}$). En la segunda, el investigador predice que, ante un docente autoritario, se observarán diferencias estadísticamente significativas entre el rendimiento de los niños que utilicen una metodología de trabajo tradicional y los que empleen un método basado en el contacto con la naturaleza ($H_0 \equiv \mu_{11} = \mu_{12}$). Por último, plantea que, ante este mismo profesor, también se producirán diferencias entre el rendimiento de los niños sometidos a una metodología tradicional y el de los que trabajen mediante un método mixto ($H_0 \equiv \mu_{11} = \mu_{13}$).

Dado que tales hipótesis se plantean *a priori*, procedemos a calcular el *rango crítico entre dos medias* aplicando el procedimiento de Bonferroni. La diferencia mínima entre dos medias (de los grupos *fg* y *hj*), para poder rechazar la hipótesis nula, viene determinada por la siguiente expresión:

$$|\bar{Y}_{fg} - \bar{Y}_{hj}| \geq \sqrt{F_{(\alpha/c, l, gl_{error})}} \sqrt{MC_{error} \sum_{i=1, j=1}^{ab} \frac{c_{ij}^2}{n_{ij}}} \quad (7.85)$$

Aplicando esta fórmula a los datos de nuestro ejemplo obtenemos:

$$|\bar{Y}_{fg} - \bar{Y}_{hj}| \geq \sqrt{F_{(0,05/3, 1, 6)}} \sqrt{58 \sum_{i=1, j=1}^6 \frac{c_{ij}^2}{n_{ij}}} \Rightarrow$$

$$\Rightarrow \sqrt{10,807} \sqrt{58 \left(\frac{1^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} + \frac{(-1)^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} \right)} \Rightarrow$$

$$\Rightarrow \sqrt{10,807} \sqrt{58 \cdot 1} = \sqrt{10,807 \cdot 58} = 25,04$$

Dado que los contrastes especificados a priori son simples, únicamente necesitamos calcular un valor de rango crítico entre medias para llevar a cabo las comparaciones que nos ocupan. En consecuencia, observamos que:

1. $|\bar{Y}_{11} - \bar{Y}_{21}| = 34 > 25,04$
2. $|\bar{Y}_{11} - \bar{Y}_{12}| = 30 > 25,04$
3. $|\bar{Y}_{11} - \bar{Y}_{13}| = 30 > 25,04$

Como la diferencia entre las medias que se comparan, en cada uno de los tres contrastes, supera el rango crítico entre dos medias obtenido mediante el procedimiento de Bonferroni, cabe concluir que $\mu_{11} \neq \mu_{21}$, $\mu_{11} \neq \mu_{12}$ y $\mu_{11} \neq \mu_{13}$.

7.3.5.2. Procedimiento DHS de Tukey

Si el investigador desea comparar de forma exhaustiva todas las posibles combinaciones entre pares de medias y dichos contrastes sean simples, el procedimiento más adecuado para tal fin es la prueba de Tukey. Dado que todas las comparaciones se realizan dos a dos, en el caso de que el diseño sea equilibrado, sólo debemos calcular una vez la distancia crítica entre dos pares de medias. Dicho rango crítico se obtiene mediante la siguiente fórmula:

$$|\bar{Y}_{fg} - \bar{Y}_{hj}| \geq \frac{q_{(z, ab, gl_{\text{error}})}}{\sqrt{2}} \sqrt{MC_{\text{error}} \sum_{i=1, j=1}^{ab} \frac{c_{ij}^2}{n_{ij}}} \quad (7.86)$$

Aplicando esta fórmula a los datos de nuestro ejemplo obtenemos:

$$|\bar{Y}_{fg} - \bar{Y}_{hj}| \geq \frac{q_{(0,05, 6, 6)}}{\sqrt{2}} \sqrt{58 \sum_{i=1, j=1}^{ab} \frac{c_{ij}^2}{n_{ij}}} \Rightarrow$$

$$\Rightarrow \frac{5,628}{\sqrt{2}} \sqrt{58 \left(\frac{1^2}{2} + \frac{(-1)^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} \right)} \Rightarrow$$

$$\Rightarrow 3,98 \sqrt{58 \cdot 1} = 30,31$$

Para simplificar la interpretación de los datos, calcularemos las distancias empíricas entre los pares de las medias implicadas en la interacción y las representaremos en una tabla (véase la Tabla 7.19).

TABLA 7.19 Diferencias entre los pares de medias de las seis condiciones experimentales que se derivan de la interacción entre los factores (Actitud del profesor y Metodología de trabajo)

Grupo	ab_{11}	ab_{12}	ab_{13}	ab_{21}	ab_{22}	ab_{23}
$ab_{11}(\bar{Y}_{11} = 42)$	0					
$ab_{12}(\bar{Y}_{12} = 12)$	30	0				
$ab_{13}(\bar{Y}_{13} = 12)$	30	0	0			
$ab_{21}(\bar{Y}_{21} = 8)$	34	4	4	0		
$ab_{22}(\bar{Y}_{22} = 20)$	22	-8	-8	-12	0	
$ab_{23}(\bar{Y}_{23} = 20)$	22	-8	-8	-12	0	0

Como cabe observar en la Tabla 7.19, aplicando este procedimiento únicamente se rechaza la hipótesis de igualdad entre los promedios para la comparación entre ab_{11} y ab_{21} . En consecuencia, podemos concluir que $\mu_{11} \neq \mu_{21}$.

Antes de abordar la prueba de Scheffé, expondremos la *propuesta realizada por Cicchetti* (1972) para maximizar la potencia del procedimiento DHS de Tukey, en los casos en los que el investigador no esté interesado en examinar todas las posibles comparaciones entre pares de medias. De hecho, con frecuencia, cuando se trabaja con un diseño $A \times B$, con JK medias marginales, el investigador únicamente desea realizar las comparaciones dos a dos entre las medias del factor A (o B) en cada nivel del factor B (o A). En tal caso, la prueba DHS de Tukey también resulta adecuada para realizar las comparaciones múltiples, pero el cálculo de su valor crítico requiere estimar un término adicional denominado «parámetro aproximado de número de medias» (J'), ya que si se utilizara como parámetro de número de medias el valor JK , se controlaría el valor de α para todas las $JK(JK - 1)/2$ posibles comparaciones dos a dos. Dicho parámetro aproximado de número de medias (J') se calcula a partir de la fórmula propuesta por Cicchetti (1972):

$$C = \frac{J'(J' - 1)}{2}$$

donde:

- J' : Parámetro aproximado de número de medias.
- C : Número de comparaciones de interés para el investigador.

Siguiendo el ejemplo propuesto por Toothaker (1991, pág. 124), imaginemos un diseño factorial intersujetos 4×3 , donde $J = 4$ y $K = 3$. En este caso, $JK = 12$. Las comparaciones entre las cuatro medias del factor A en cada nivel del factor B darían lugar a $J(J - 1)/2 = 4(4 - 1)/2 = 6$ comparaciones en cada uno de los tres niveles del factor B . Así, para los tres niveles de este factor, existirían $6K = 18$ comparaciones de interés. No obstante, existen $JK(JK - 1)/2 = 12(12 - 1)/2 = 66$ posibles comparaciones dos a dos, entre las 12 medias marginales. Como puede constatarse, las 18 comparaciones que le interesan al investigador constituyen un subconjunto del total de las 66 posibles comparaciones. Por ello, con el fin de obtener el parámetro de número de medias necesario para el cálculo del valor

crítico de la prueba DHS de Tukey, el hecho de escoger un valor $JK = 12$ supondría la utilización de pruebas muy conservadoras, mientras que la elección de un valor de $J = 4$ implicaría la utilización de pruebas muy liberales. Con el fin de maximizar la potencia de la prueba DHS de Tukey, debemos calcular el *parámetro aproximado de número de medias* a partir de la ecuación de Cicchetti. Aplicando dicha ecuación obtenemos la siguiente igualdad: $18 = J'(J' - 1)/2$, de modo que $36 = J'(J' - 1)$. Así, los posibles valores de J' son:

$$6(6 - 1)/2 = 30/2 = 15 \quad \text{y} \quad 7(7 - 1)/2 = 42/2 = 21$$

De entre los dos posibles valores, a saber, $J' = 6$ o $J' = 7$, este último es el valor de J' más cercano a una solución de 18 comparaciones de interés, que no implica la utilización de una prueba liberal⁶. El valor de q (distribución de rango estudentizado) para los parámetros $\alpha = 0,05$, $J' = 7$ y $gl_{\text{error}} = 60$, es de $q = 4,31$ (ver Tablas 5, 6 y 7 en el Anexo A). Si se desea utilizar la prueba DHS de Tukey como estrategia de comparación múltiple, debemos comparar el valor empírico obtenido en dicha prueba con un valor crítico de $q/\sqrt{2}$. En cualquier caso, rechazaremos la hipótesis nula de igualdad entre las medias comparadas cuando el valor empírico supere el valor crítico.

7.3.5.3. Prueba de Scheffé

Esta prueba es válida para cualquier circunstancia, tanto si se realizan comparaciones *a priori* como *a posteriori* y tanto si tales comparaciones son *simples* como *complejas*. La distancia crítica entre pares de medias se obtiene mediante la siguiente fórmula:

$$|\bar{Y}_{fg} - \bar{Y}_{hj}| \geq \sqrt{(ab - 1)F_{(\alpha, ab - 1, gl_{\text{error}})}} \sqrt{MC_{\text{error}} \sum_{i=1, j=1}^{ab} \frac{c_{ij}^2}{n_{ij}}} \quad (7.87)$$

Aplicando esta fórmula a los datos de nuestro ejemplo obtenemos:

$$\begin{aligned} |\bar{Y}_{fg} - \bar{Y}_{hj}| &\geq \sqrt{(6 - 1)F_{(0,05, 6 - 1, 6)}} \sqrt{58 \sum_{i=1, j=1}^{ab} \frac{c_{ij}^2}{n_{ij}}} \Rightarrow \\ &\Rightarrow \sqrt{5 \cdot 4,387} \sqrt{58 \left(\frac{1^2}{2} + \frac{(-1)^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} + \frac{0^2}{2} \right)} \Rightarrow \\ &\Rightarrow 4,68 \cdot 7,61 = 35,64 \end{aligned}$$

Como ya apuntamos al abordar las comparaciones múltiples entre medias, en el diseño multigrupos aleatorios (véase el Epígrafe 6.2.2.4 del Capítulo 6), aunque el procedimiento

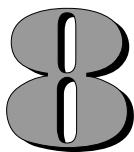
⁶ La elección de un valor $J' = 6$ da lugar a una solución de $C = 15$ comparaciones de interés, valor que es inferior al número de comparaciones de interés del investigador ($C = 18$). La elección de este valor, como parámetro aproximado de número de medias, implicaría la utilización de un valor crítico liberal de la prueba DHS de Tukey, es decir, tendríamos una alta probabilidad de cometer un error de tipo I (aunque la potencia de la prueba sería mayor). Por el contrario, la elección de un valor $J' = 7$ da lugar a una solución de $C = 21$ comparaciones de interés, valor que es superior al número de comparaciones de interés del investigador ($C = 18$). La elección de este valor implicaría la utilización de un valor crítico algo más conservador que, aunque disminuye la potencia de la prueba, controla adecuadamente el error de tipo I.

de Scheffé permite controlar el error de tipo I sin restricciones con respecto al número de comparaciones que se efectúan, esta estrategia es altamente conservadora. Ello hace que, en nuestro caso, la hipótesis nula de igualdad entre los promedios no pueda rechazarse en ninguna de las comparaciones derivadas de la interacción entre los factores *A* y *B*.

7.3.5.4. Análisis de datos mediante el paquete estadístico SPSS 10.0

Para realizar las pruebas de comparaciones múltiples, han de seguirse las pautas descritas en el capítulo anterior, tanto en lo que respecta a la elección del procedimiento (Epígrafe 6.2.2.4) como a su cálculo mediante el paquete estadístico SPSS 10.0 (Epígrafe 6.2.2.5).

Como ya se ha indicado anteriormente, en el caso de que la interacción resultara estadísticamente significativa y de que estuviéramos interesados en realizar todas las posibles comparaciones dos a dos, deberíamos llevar a cabo una transformación en las variables, con el fin de crear una nueva variable cuyos niveles correspondiesen a los tratamientos resultantes de la combinación entre los niveles de los factores *A* y *B*. Dicho de otro modo, tras realizar la transformación, la combinación entre los niveles del factor *A* y del factor *B* genera el conjunto de categorías de la nueva variable que se crea mediante el programa SPSS. Para llevar a cabo dicha transformación, remitimos al lector al subapartado «Comparaciones múltiples» del Epígrafe 7.3.2.4.



DISEÑOS EXPERIMENTALES QUE REDUCEN LA VARIANZA DE ERROR

Los diseños experimentales que se han abordado hasta el momento utilizan la *aleatorización* como única técnica de control. En consecuencia, generan una cantidad considerable de varianza de error, lo que disminuye la sensibilidad o la potencia del diseño. Existen diversas estrategias metodológicas, tanto de naturaleza experimental como estadística, para optimizar el diseño o, en otras palabras, para reducir la varianza de error y conseguir una mayor precisión en la estimación de los efectos. Entre las estrategias de carácter experimental cabe destacar el *emparejamiento*, el *bloqueo*, el *anidamiento* y la *utilización del sujeto como control propio*. En lo que respecta a los procedimientos de control estadístico, el método más utilizado en el ámbito de las ciencias del comportamiento es el *análisis de la covarianza*.

Dado que la **técnica de emparejamiento** se emplea únicamente con diseños de dos grupos y que su lógica es similar a la que subyace a la *técnica de bloqueo*, no consideramos necesario profundizar en los diseños asociados a dicho procedimiento de control o **diseños de dos grupos emparejados**. No obstante, describiremos brevemente en qué consiste tal técnica. La técnica de emparejamiento consiste en igualar a dos grupos de sujetos con respecto a una variable específica. Así, partiendo de una muestra obtenida mediante un procedimiento de selección aleatoria, se forman pares de sujetos con aquellos miembros de la muestra que obtienen la misma puntuación (o una puntuación similar) en una determinada variable criterio, denominada *variable de emparejamiento*. Esta variable es una variable extraña que, a juicio del experimentador, influye de forma decisiva en la variable dependiente. Tras emparejar a los sujetos en función de dicha variable, se asigna aleatoriamente un miembro de cada par a cada uno de los grupos de tratamiento. De este modo, se controla una posible fuente de varianza de error y se incrementa la potencia estadística del diseño. En relación con los análisis estadísticos asociados a los diseños de dos grupos emparejados, cabe señalar, en términos generales, que cuando los datos son de naturaleza paramétrica la **prueba de la hipótesis** se lleva a cabo aplicando la *t de Student* para muestras independientes. En caso de que los datos sean no paramétricos, se deben utilizar *pruebas estadísticas basadas en rangos o en frecuencias*.

8.1. DISEÑOS CON UN FACTOR DE BLOQUEO: DISEÑOS DE BLOQUES ALEATORIOS

8.1.1. Características generales del diseño de bloques aleatorios

El principio básico de la *técnica de bloqueo* consiste en formar bloques homogéneos de sujetos o de unidades experimentales en función de las puntuaciones presentadas por dichos sujetos en una variable extraña relevante. De esta forma, los sujetos pertenecientes a cada bloque proceden de un determinado estrato de la población de origen y presentan valores similares en alguna característica psicológica, biológica o social. Entre las características que se tienen en cuenta con mayor frecuencia, Spector (1993) destaca el sexo, la inteligencia, el nivel socioeconómico y el lugar de residencia. Esta variable de clasificación o variable criterio, que sirve para agrupar a los sujetos en grupos homogéneos, recibe el nombre de *variable de bloqueo*. Tras la formación de grupos de sujetos equivalentes, en función de dicha variable, se asignan aleatoriamente los sujetos de cada uno de los bloques a las diferentes condiciones de tratamiento. Por tanto, la técnica de bloqueo supone aplicar una restricción al azar, ya que los sujetos son asignados aleatoriamente a los grupos a partir de los bloques. Por otra parte, dado que dentro de cada bloque se forma una cantidad de grupos de sujetos igual al número de tratamientos experimentales, cada bloque puede considerarse como una réplica completa del experimento. Es importante señalar que la puntuación del sujeto en la variable de bloqueo sirve, fundamentalmente, como criterio de adscripción a una variable categórica con valores fijos.

Los diseños asociados a esta técnica de control reciben el nombre de *diseños de bloques aleatorios* o *diseños parcialmente aleatorizados*. La principal ventaja de estos diseños, con respecto a los diseños completamente aleatorios, radica en que permiten estimar, con mayor precisión, los efectos que ejerce(n) la(s) variable(s) independiente(s) sobre la(s) variable(s) dependiente(s). Ello es así porque la creación de grupos de sujetos equivalentes u homogéneos permite eliminar una potencial fuente de variabilidad extraña relacionada con las propias unidades experimentales y, en consecuencia, posibilita reducir la varianza de error. En definitiva, los diseños de bloques aleatorios son estructuras de investigación en las que se aplica la técnica de control experimental denominada técnica de bloqueo, con el objetivo de reducir el término residual del análisis de la varianza mediante el control de una variable extraña relacionada con la variable dependiente.

En principio, por su propia naturaleza, el diseño de bloques no asume interacción entre la variable manipulada y la de bloqueo, ya que su finalidad no radica en examinar los efectos de la variable de bloqueo como si se tratara de un factor explicativo, sino en reducir la varianza de error. Por la propia concepción de la técnica de bloqueo, se supone que la variable de bloqueo constituye una fuente de variación estrechamente relacionada con la variable dependiente, pero ajena a los intereses del investigador. De hecho, la razón principal por la que se incluye en la ecuación estructural del diseño consiste en evitar que forme parte del componente residual del modelo y que reduzca la sensibilidad del diseño. Sin embargo, en algunos casos, el factor manipulado y el factor de bloqueo interaccionan entre sí, de manera que la variable de bloqueo se convierte en un factor explicativo que condiciona la influencia que ejerce el factor manipulado sobre la variable dependiente. Por otra parte, para que la técnica de bloqueo sea realmente efectiva, la variable de bloqueo tiene que estar directamente relacionada con la variable dependiente. De esta forma, si se rechaza la hipótesis nula para el factor de bloqueo, se consigue eliminar del término de error la proporción de varianza explicada por dicho factor y cumplir con la finalidad básica del diseño, a saber, aumentar la precisión en la estimación de los efectos experimentales.

8.1.2. Análisis factorial de la varianza para el diseño de bloques aleatorios

El *análisis factorial de la varianza* es el modelo analítico que se utiliza, habitualmente, para llevar a cabo la **prueba de la hipótesis** en los diseños de bloques aleatorios. No obstante, cabe distinguir dos estructuras de diseño, en función de las cuales el modelo matemático subyacente a la predicción que se realiza bajo la hipótesis alternativa responde a expresiones diferentes. Tales estructuras son el *diseño de un solo sujeto por casilla o por combinación entre tratamiento y bloque* y el *diseño con más de un sujeto por casilla o por combinación entre tratamiento y bloque*. En el primer tipo de diseño, si se produce una interacción entre la variable de bloqueo y la variable manipulada, tal interacción coincide con el error experimental y, por tanto, se utiliza como término de contrastación para verificar el efecto de los tratamientos. En el segundo modelo de diseño, por el contrario, es posible estimar el término de interacción entre la variable de bloqueo y la manipulada y el término de error de forma independiente. En los dos epígrafes siguientes abordamos el análisis factorial de la varianza para cada uno de estos dos modelos de diseño.

8.1.2.1. Análisis factorial de la varianza para el diseño de un solo sujeto por casilla o por combinación entre tratamiento y bloque

• Modelo general de análisis

Suponiendo una **estructura básica de un solo sujeto por casilla**, la ecuación estructural, asociada a la predicción que se realiza bajo la hipótesis alternativa, puede ajustarse a dos modelos diferentes: el *modelo aditivo* y el *modelo no aditivo* (Arnaud, 1986).

El *modelo estructural aditivo* responde a la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \beta_k + \varepsilon_{ijk} \quad (8.1)$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable A o variable de tratamiento y el k -ésimo nivel de la variable B o variable de bloqueo.

μ = Media general de la variable dependiente o término que recoge el efecto de todos los factores constantes.

α_j = Efecto debido a la administración del j -ésimo nivel de la variable de tratamiento.

β_k = Efecto debido a la administración del k -ésimo nivel de la variable de bloqueo.

ε_{ijk} = Término de error o componente aleatorio del modelo. Se asume que $\varepsilon_{ijk} \sim NID(0, \sigma^2)$.

El modelo aditivo se basa en el supuesto de que no existe interacción entre la variable de tratamiento y la variable de bloqueo.

Sin embargo, es posible que los datos del experimento no se ajusten al modelo de aditividad y que, en consecuencia, resulte más adecuado representarlos mediante un modelo más completo. Así, en el supuesto de que el valor esperado del término residual no fuera cero, sino que cada residual tuviera su propio valor esperado, el término ε_{ijk} debería incluir un componente de no aditividad: γ_{ijk} . De esta forma:

$$\varepsilon_{ijk} = \gamma_{ijk} + \eta_{ijk}$$

donde:

γ_{ijk} = Componente no aditivo que recoge la asunción de que el efecto de la variable de bloqueo (β_k) se combina con el efecto de la variable de tratamiento (α_j) siguiendo, no una relación lineal, sino multiplicativa.

η_{ijk} = Término residual del modelo, para el que se asume un valor esperado igual a cero.

En definitiva, cuando no se puede asumir que la interacción entre la variable manipulada y la variable de bloqueo es nula, el modelo debe incluir un componente de no aditividad que represente dicha interacción, ya que en caso de no tener en cuenta este componente, el cálculo de las razones F para los factores principales resultaría negativamente sesgado. Tukey (1949) ha desarrollado un procedimiento para calcular el componente de no aditividad y para comprobar si ejerce una influencia, estadísticamente significativa, sobre la variable de respuesta.

Teniendo en cuenta dicho componente, el *modelo estructural no aditivo* propio del diseño de bloques aleatorios de un solo sujeto por casilla responde a la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \beta_k + \gamma_{ijk} + \eta_{ijk} \quad (8.2)$$

cuyos términos ya han sido descritos en este mismo epígrafe.

• Ejemplo práctico

Supongamos que, en el ámbito de la psicología educativa, realizamos una investigación para examinar la influencia que ejerce la *metodología de trabajo empleada en el aula* (factor A) sobre el *rendimiento presentado por un grupo de niños en la asignatura de Ciencias Naturales*. No obstante, se considera que una posible variable extraña, capaz de contaminar los resultados del estudio, es el *nivel de motivación de los sujetos* (factor B). Con el objeto de controlar dicha variable se utiliza un diseño de bloques aleatorios. Se dividen los niños en cuatro bloques, de tres sujetos cada uno, en función de las puntuaciones que obtienen en una prueba destinada a medir su nivel de motivación con respecto al aprendizaje, a saber: (b_1) *nivel de motivación bajo*, (b_2) *nivel de motivación medio-bajo*, (b_3) *nivel de motivación medio-alto* y (b_4) *nivel de motivación alto*. A continuación, a cada uno de los sujetos que configuran cada bloque se le somete a una de las tres siguientes metodologías de trabajo: (a_1) *metodología basada en medios audiovisuales*, (a_2) *metodología basada en el contacto directo con la naturaleza* y (a_3) *metodología tradicional*. En la Tabla 8.1 pueden observarse los resultados obtenidos en la investigación.

En el presente modelo de diseño únicamente utilizaremos el primero de los procedimientos que hemos empleado para el cálculo de las sumas de cuadrados en los diseños intergrupos aleatorios (véase el Epígrafe 7.3.2.3) y, a partir de tales sumas cuadráticas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la *ecuación estructural ajustada al modelo aditivo*. Posteriormente desarrollaremos la prueba que permite verificar la hipótesis de nulidad del componente de no aditividad o la *prueba de Tukey*, a fin de dilucidar si la interacción entre la variable manipulada y la variable de bloqueo debe ser incluida en el modelo. Dado que normalmente se trabaja con diseños de bloques aleatorios que tienen más de un sujeto por casilla, pospondremos el cálculo de las sumas de cuadrados, mediante diferentes procedimientos, para la sección en la que se aborda el ANOVA asociado a dicho modelo de diseño.

TABLA 8.1 Matriz de datos del experimento

		A (Metodología de trabajo)			Sumatorios y medias marginales
		a_1 (Audiovisual)	a_2 (Contacto con la naturaleza)	a_3 (Tradicional)	
B (Nivel de motivación)	b_1 (Bajo)	4	14	2	$\Sigma Y_{.1} = 20$ $\bar{Y}_{.1} = 6,66$
	b_2 (Medio-bajo)	7	12	4	$\Sigma Y_{.2} = 23$ $\bar{Y}_{.2} = 7,66$
	b_3 (Medio-alto)	8	15	5	$\Sigma Y_{.3} = 28$ $\bar{Y}_{.3} = 9,33$
	b_4 (Alto)	12	19	5	$\Sigma Y_{.4} = 36$ $\bar{Y}_{.4} = 12$
Sumatorios y medias marginales		$\Sigma Y_{1.} = 31$ $\bar{Y}_{1.} = 7,75$	$\Sigma Y_{2.} = 60$ $\bar{Y}_{2.} = 15$	$\Sigma Y_{3.} = 16$ $\bar{Y}_{3.} = 4$	$\Sigma Y_{..} = 107$ $\bar{Y}_{..} = 8,92$

Procedamos a desarrollar el análisis factorial de la varianza, para el diseño de bloques aleatorios con un solo sujeto por casilla, tomando como referencia la matriz de datos de la Tabla 8.1.

Fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas:

- Efecto del factor de tratamiento o factor A:

$$SCA = \left[\frac{1}{b} \sum_j \left(\sum_k Y_{jk} \right)^2 \right] - C \quad (8.3)$$

- Efecto del factor de bloqueo o factor B:

$$SCB = \left[\frac{1}{a} \sum_k \left(\sum_j Y_{jk} \right)^2 \right] - C \quad (8.4)$$

- Variabilidad total:

$$SCT = \sum_j \sum_k Y_{jk}^2 - C \quad (8.5)$$

- Variabilidad residual o del error:

$$SCR = SCT - SCA - SCB \quad (8.6)$$

El componente C de las fórmulas precedentes se calcula mediante la siguiente expresión:

$$C = \frac{1}{ab} \left(\sum_j \sum_k Y_{jk} \right)^2 \quad (8.7)$$

Tras obtener el valor de C , llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{3 \cdot 4} (4 + 7 + 8 + 12 + 14 + 12 + 15 + 19 + 2 + 4 + 5 + 5)^2 = \frac{(107)^2}{12} = 954,08$$

Una vez obtenido este valor, procedemos a calcular la SCA , la SCB , la SCT y la SCR , respectivamente:

$$SCA = \frac{1}{4} [(31)^2 + (60)^2 + (16)^2] - 954,08 = 1.204,25 - 954,08 = 250,17$$

$$SCB = \frac{1}{4} [(20)^2 + (23)^2 + (28)^2 + (36)^2] - 954,08 = 1.003 - 954,08 = 48,92$$

$$SCT = [(4)^2 + (7)^2 + (8)^2 + (12)^2 + (14)^2 + (12)^2 + (15)^2 + (19)^2 + (2)^2 + (4)^2 + (5)^2 + (5)^2] - 954,08 = 1.269 - 954,08 = 314,92$$

$$SCR = 314,92 - 250,17 - 48,92 = 15,83$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 8.2 Análisis factorial de la varianza para el diseño de bloques aleatorios con un solo sujeto por casilla bajo el modelo de aditividad: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Entre tratamientos (factor A)	$SCA = 250,17$	$a - 1 = 2$	$MCA = 125,08$	$F_A = 47,38$
Entre bloques (factor B)	$SCB = 48,92$	$b - 1 = 3$	$MCB = 16,31$	$F_B = 6,18$
Residual o del error (tratamientos $\times$ bloques)	$SCR = 15,83$	$(a - 1)(b - 1) = 6$	$MCR = 2,64$	
Total	$SCT = 314,92$	$ab - 1 = 11$		

Tras consultar las *tablas de los valores críticos de la distribución F* con un nivel de confianza del 95% ($\alpha = 0,05$) y trabajando con una hipótesis de una cola, podemos concluir que la *metodología empleada en el aula* (factor A) ejerce una influencia estadísticamente significativa sobre el *rendimiento de los niños en la asignatura de Ciencias Naturales* ($F_{0,95; 2,6} = 5,14 < F_{\text{obs.}} = 47,38$).

A su vez, la *variable de bloqueo* (factor B) también afecta significativamente al rendimiento ($F_{0,95; 3,6} = 4,76 < F_{\text{obs.}} = 6,18$).

PRUEBA DEL COMPONENTE DE NO ADITIVIDAD: PRUEBA DE TUKEY

Los cálculos necesarios para aplicar la prueba de Tukey se representan en la Tabla 8.3.

TABLA 8.3 Cálculos necesarios para aplicar la prueba de no aditividad de 1 g.l. de Tukey

	a_1	a_2	a_3	$\bar{Y}_{.k}$	$(\bar{Y}_{.k} - \bar{Y}_{..})$	$(\bar{Y}_{.k} - \bar{Y}_{..})^2$	$\sum_j Y_{jk} (\bar{Y}_{j.} - \bar{Y}_{..})$ (productos cruzados)
b_1	4	14	2	6,66	-2,26	5,11	$4(-1,17) + 14(6,08) + 2(-4,92) = 70,6$
b_2	7	12	4	7,66	-1,26	1,59	$7(-1,17) + 12(6,08) + 4(-4,92) = 45,09$
b_3	8	15	5	9,33	0,41	0,17	$8(-1,17) + 15(6,08) + 5(-4,92) = 57,24$
b_4	12	19	5	12	3,08	9,49	$12(-1,17) + 19(6,08) + 5(-4,92) = 76,88$
$\bar{Y}_{j.}$	7,75	15	4	$\bar{Y}_{..} = 8,92$		$\sum_k (\bar{Y}_{.k} - \bar{Y}_{..})^2 = 16,36$	
$(\bar{Y}_{j.} - \bar{Y}_{..})$	-1,17	6,08	-4,92				
$(\bar{Y}_{j.} - \bar{Y}_{..})^2$	1,37	36,97	24,21				
$\sum_j (\bar{Y}_{j.} - \bar{Y}_{..})^2 = 62,55$							

Tras realizar los cálculos de la Tabla 8.3, podemos hallar la suma de cuadrados de la no aditividad. El numerador de dicha suma cuadrática se obtiene mediante la siguiente expresión:

$$\begin{aligned} \text{Numerador } SC_{\text{no adit.}} &= \left[\sum_k (\bar{Y}_{.k} - \bar{Y}_{..}) \sum_j Y_{jk} (\bar{Y}_{j.} - \bar{Y}_{..}) \right]^2 = \\ &= [(-2,26)(70,6) + (-1,26)(45,09) + (0,41)(57,24) + (3,08)(76,88)]^2 = \quad (8.8) \\ &= [(-159,556) + (-56,81) + (23,47) + (236,79)]^2 = (43,894)^2 = 1.926,68 \end{aligned}$$

El denominador se obtiene aplicando la siguiente fórmula:

$$\begin{aligned} \text{Denominador } SC_{\text{no adit.}} &= \left[\sum_k (\bar{Y}_{.k} - \bar{Y}_{..})^2 \right] \left[\sum_j (\bar{Y}_{j.} - \bar{Y}_{..})^2 \right] = \\ &= (16,36)(62,55) = 1.023,32 \quad (8.9) \end{aligned}$$

Por tanto, la *suma de cuadrados de la no aditividad* es:

$$SC_{\text{no adit.}} = \frac{1.926,68}{1.023,32} = 1,88 \quad (8.10)$$

Para calcular la *suma cuadrática residual de la no aditividad* o el término de contraste necesario para la prueba de la hipótesis del componente de no aditividad, debemos sustraer

la suma cuadrática de la no aditividad, de la suma cuadrática residual del ANOVA. Así, obtenemos:

$$SC_{\text{residual (no adit.)}} = SC_{\text{residual}} - SC_{\text{no adit.}} = 15,83 - 1,88 = 13,95 \tag{8.11}$$

Esta expresión es equivalente a:

$$\begin{aligned} SC_{\text{residual (no adit.)}} &= SCT - SCA - SCB - SC_{\text{no adit.}} = \\ &= 314,92 - 250,17 - 48,92 - 1,88 = 13,95 \end{aligned} \tag{8.12}$$

En la Tabla 8.4 presentamos el cuadro resumen del análisis de la varianza para el diseño de bloques aleatorios con un solo sujeto por casilla bajo el modelo de no aditividad.

TABLA 8.4 Análisis factorial de la varianza para el diseño de bloques aleatorios con un solo sujeto por casilla bajo el modelo de no aditividad: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Entre tratamientos (factor A)	SCA = 250,17	2	MCA = 125,08	$F_{\text{no adit.}} = 0,67$
Entre bloques (factor B)	SCB = 48,92	3	MCB = 16,31	
Componente de no aditividad	SC _{no adit.} = 1,88	1	MC _{no adit.} = 1,88	
Residual o del error	SCR _{no adit.} = 13,95	5	MCR _{no adit.} = 2,79	
Total	SCT = 314,92	11		

Para determinar si el componente de no aditividad es o no estadísticamente significativo, debemos recurrir a las *tablas de los valores críticos de la distribución F*. Dado que con un nivel de confianza del 95 % ($\alpha = 0,05$) observamos que $F_{0,95; 1,5} = 6,61 > F_{\text{obser.}} = 0,67$, cabe concluir que dicho componente no es estadísticamente significativo. En consecuencia, los datos del experimento se ajustan al modelo de aditividad, es decir, al modelo en el que se asume que no existe interacción entre la variable de tratamiento y la variable de bloqueo.

8.1.2.2. Análisis factorial de la varianza para el diseño con más de un sujeto por casilla o por combinación entre tratamiento y bloque

• **Modelo general de análisis**

Como se ha señalado anteriormente, el modelo de diseño de bloques aleatorios que se utiliza habitualmente consta de dos o más sujetos por combinación entre tratamiento y bloque. La ecuación estructural del ANOVA asociada a la predicción que se realiza bajo la hipótesis alternativa, cuando se utiliza un **diseño de bloques aleatorios con más de un sujeto por casilla**, se expresa mediante la combinación lineal de los siguientes parámetros:

$$y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \tag{8.13}$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable A o variable de tratamiento y el k -ésimo nivel de la variable B o variable de bloqueo.

μ = Media general de la variable dependiente.

α_j = Efecto debido a la administración del j -ésimo nivel de la variable de tratamiento.

β_k = Efecto debido a la administración del k -ésimo nivel de la variable de bloqueo.

$(\alpha\beta)_{jk}$ = Efecto debido a la interacción entre el j -ésimo nivel de la variable de tratamiento y el k -ésimo nivel de la variable de bloqueo.

ε_{ijk} = Término de error o componente aleatorio del modelo. Como en todo modelo estructural de análisis de la varianza, se asume que los errores son independientes y tienen una distribución normal con media igual a cero y varianza σ^2 .

Dado que en este tipo de diseño disponemos de más de un sujeto por casilla, se puede calcular el componente residual del modelo y, por tanto, la interacción no se confunde con el término de error. Esta estructura de diseño se utiliza con mucha mayor frecuencia que la que consta de un solo sujeto por casilla y su solución analítica es exactamente igual a la del diseño factorial de dos factores al azar. En ambos casos el análisis factorial de la varianza parte de la misma descomposición de las fuentes de variación y llega a la estimación de las razones F para los mismos parámetros. No obstante, no debemos olvidar que, mientras en el diseño factorial de dos factores al azar las dos variables independientes son experimentales, en el diseño de bloques aleatorios una de las variables es de clasificación.

• Ejemplo práctico

Consideremos la misma investigación que hemos tomado como referencia para desarrollar el ANOVA en el diseño de bloques aleatorios con un solo sujeto por casilla, pero suponiendo que asignamos dos sujetos a cada casilla. Así pues, disponemos de un total de seis sujetos por bloque y ocho sujetos por condición experimental. En la Tabla 8.5 se muestra la puntuación media obtenida por cada uno de los sujetos en una batería de pruebas destinada a medir el rendimiento en la asignatura de Ciencias Naturales.

La Tabla 8.6 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan las *categorías del factor de bloqueo* o *factor B* (en el ejemplo: b_1, b_2, b_3 y b_4) y en las columnas, los *niveles del factor A* (en el ejemplo: a_1, a_2 y a_3). Dado que cada combinación entre tratamiento y bloque consta de dos sujetos, los subíndices correspondientes al modelo de diseño de nuestro ejemplo son $i = 1, 2$ (sujetos dentro de cada casilla), $j = 1, 2, 3$ (niveles del factor de tratamiento) y $k = 1, 2, 3, 4$ (niveles del factor de bloqueo).

Examinemos con más detalle la estructura de la tabla. Cada una de las casillas de la tabla representa un *grupo de sujetos sometido a una determinada combinación entre el factor de tratamiento y el factor de bloqueo*. En la parte derecha (en la columna $T^2_{..k}$) situamos los *cuadrados de los sumatorios* correspondientes a las filas ($k = 1, 2, \dots, b$) y en la parte inferior de la tabla (en la fila $T^2_{.j.}$) los *cuadrados de los sumatorios* correspondientes a las columnas ($j = 1, 2, \dots, a$). Como cabe apreciar, los primeros corresponden a los *niveles del factor B* y los segundos a las *categorías del factor A*. En los márgenes derecho e inferior hemos situado las *medias aritméticas* correspondientes a las filas y a las columnas, respectivamente.

TABLA 8.5 Matriz de datos del experimento

		A (Metodología de trabajo)			Cuadrados de sumatorios y medias marginales
		a_1 (Audiovisual)	a_2 (Contacto con la naturaleza)	a_3 (Tradicional)	
B (Nivel de motivación)	b_1 (Bajo)	4 3	14 12	2 1	$\Sigma Y^2_{..1} = 1.296$ $\bar{Y}_{..1} = 6$
	b_2 (Medio-bajo)	7 6	12 10	4 2	$\Sigma Y^2_{..2} = 1.681$ $\bar{Y}_{..2} = 6,83$
	b_3 (Medio-alto)	8 3	15 14	5 4	$\Sigma Y^2_{..3} = 2.401$ $\bar{Y}_{..3} = 8,16$
	b_4 (Alto)	12 11	19 13	5 3	$\Sigma Y^2_{..4} = 3.969$ $\bar{Y}_{..4} = 10,5$
Cuadrados de sumatorios y medias marginales		$\Sigma Y^2_{.1.} = 2.916$ $\bar{Y}_{.1.} = 6,75$	$\Sigma Y^2_{.2.} = 11.881$ $\bar{Y}_{.2.} = 13,62$	$\Sigma Y^2_{.3.} = 676$ $\bar{Y}_{.3.} = 3,25$	$\Sigma Y^2_{...} = 35.721$ $\bar{Y}_{...} = 7,87$

TABLA 8.6 Datos correspondientes a un diseño de bloques aleatorios con un factor de tratamiento y un factor de bloqueo y con más de un sujeto por casilla: modelo general

Categorías		Factor de tratamiento (A)					Cuadrados de sumatorios marginales $T^2_{..k}$	Medias marginales k
		a_1	...	a_j	...	a_a		
Factor de bloqueo (B)	b_1	Y_{111} Y_{i11} Y_{n11}	...	Y_{1j1} Y_{ij1} Y_{nj1}	...	Y_{1a1} Y_{ia1} Y_{na1}	$T^2_{..1}$	$\bar{Y}_{..1}$
	...		...		...		...	
	b_k	Y_{11k} Y_{i1k} Y_{n1k}	...	Y_{1jk} Y_{ijk} Y_{nj1}	...	Y_{1ak} Y_{iak} Y_{nak}	$T^2_{..k}$	$\bar{Y}_{..k}$
	...		...		...		...	
	b_b	Y_{11b} Y_{i1b} Y_{n1b}	...	Y_{1jb} Y_{ijb} Y_{njb}	...	Y_{1ab} Y_{iab} Y_{nab}	$T^2_{..b}$	$\bar{Y}_{..b}$
Cuadrados de sumatorios marginales $T^2_{.j.}$		$T^2_{.1.}$		$T^2_{.j.}$		$T^2_{.a.}$	$T^2_{...}$	
Medias marginales j		$\bar{Y}_{.1.}$		$\bar{Y}_{.j.}$		$\bar{Y}_{.a.}$		$\bar{Y}_{...}$

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño y, suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 8. 5.

• Desarrollo del análisis factorial de la varianza

En el presente diseño, seguiremos la misma secuencia empleada en el caso de los diseños factoriales completamente aleatorios para el cálculo de las sumas de cuadrados, de las varianzas y de las razones F asociadas a los distintos parámetros de la ecuación estructural del ANOVA.

Procedimiento 1

En la Tabla 8.7 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas.

TABLA 8.7 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño de bloques aleatorios con un factor de tratamiento y un factor de bloqueo y con más de un sujeto por casilla

Entre tratamientos (factor A)	$SCA = \left[\frac{1}{bn} \sum_j \left(\sum_i Y_{ijk} \right)^2 \right] - C$	(8.14)
Entre bloques (factor B)	$SCB = \left[\frac{1}{an} \sum_k \left(\sum_j Y_{ijk} \right)^2 \right] - C$	(8.15)
Entre grupos	$SCEG = \left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C$	(8.16)
Interacción $A \times B$	$SCAB = SCEG - SCA - SCB$	(8.17)
Residual o del error	$SCR = SCT - (SCA + SCB + SCAB)$	(8.18)
Variabilidad total	$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$	(8.19)
C	$C = \frac{1}{abn} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2$	(8.20)

Tras obtener el valor de C , llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{abn} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2$$

Término a término:

$$\frac{1}{abn} = \frac{1}{3 \cdot 4 \cdot 2} = \frac{1}{24}$$

$$\left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2 = (4 + 3 + 7 + 6 + 8 + 3 + 12 + 11 + 14 + 12 + 12 + 10 + 15 + 14 + 19 + 13 + 2 + 1 + 4 + 2 + 5 + 4 + 5 + 3)^2 = (189)^2$$

Por tanto:

$$C = \frac{(189)^2}{24} = 1.488,375$$

Una vez obtenido este valor, procedemos a calcular la *SCA*, la *SCB*, la *SCEG*, la *SCR* y la variabilidad correspondiente a la suma de todas las anteriores, a saber, la *SCT*. Comencemos con el cálculo de la suma de cuadrados entre tratamientos o *SCA*. Desarrollando la Fórmula (8.14) obtenemos:

$$\begin{aligned} SCA &= \left[\frac{1}{bn} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C \\ SCA &= \left[\frac{1}{4 \cdot 2} [(4 + 3 + 7 + 6 + 8 + 3 + 12 + 11)^2 + \right. \\ &\quad \left. + (14 + 12 + 12 + 10 + 15 + 14 + 19 + 13)^2 + (2 + 1 + 4 + 2 + 5 + 4 + 5 + 3)^2] \right] - C \\ SCA &= \left[\frac{1}{8} [(54)^2 + (109)^2 + (26)^2] \right] - C = 1.934,125 - 1.488,375 = 445,75 \end{aligned}$$

Para calcular la variabilidad entre bloques o asociada al factor *B*, aplicamos la Fórmula (8.15):

$$\begin{aligned} SCB &= \left[\frac{1}{an} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C \\ SCB &= \left[\frac{1}{3 \cdot 2} [(4 + 3 + 14 + 12 + 2 + 1)^2 + (7 + 6 + 12 + 10 + 4 + 2)^2 + \right. \\ &\quad \left. + (8 + 3 + 15 + 14 + 5 + 4)^2 + (11 + 12 + 19 + 13 + 5 + 3)^2] \right] - C \\ SCB &= \frac{1}{6} [(36)^2 + (41)^2 + (49)^2 + (63)^2] - C = 1.557,833 - 1.488,375 = 69,46 \end{aligned}$$

A partir de la Fórmula (8.16) obtenemos la variabilidad entre grupos. Cabe señalar que la *SCEG* también puede calcularse mediante la suma de la variabilidad asociada al factor *A*, la variabilidad asociada al factor *B* y la debida a la interacción entre ambos factores, a saber, $SCEG = SCA + SCB + SCAB$.

$$SCEG = \left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C$$

$$\sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 =$$

$$= \left[(4+3)^2 + (7+6)^2 + (8+3)^2 + (12+11)^2 + (14+12)^2 + (12+10)^2 + \right.$$

$$\left. + (15+14)^2 + (19+13)^2 + (2+1)^2 + (4+2)^2 + (5+4)^2 + (5+3)^2 \right]$$

$$SCEG = 2.041,5 - 1.488,375 = 553,125$$

La variabilidad asociada a la interacción entre el factor *A* y el factor *B* se calcula aplicando la fórmula (8.17):

$$SCAB = SCEG - SCA - SCB$$

$$SCAB = 553,125 - 445,75 - 69,46 = 37,916$$

Mediante la Fórmula (8.19) obtenemos la variabilidad total. Ya se ha señalado, en múltiples ocasiones, que la *SCT* constituye la suma de las restantes variabilidades, es decir, $SCT = SCA + SCB + SCAB + SCR$.

$$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$$

$$\sum_i \sum_j \sum_k Y_{ijk}^2 = \left[(4)^2 + (3)^2 + (7)^2 + (6)^2 + (8)^2 + (3)^2 + (12)^2 + (11)^2 + (14)^2 + (12)^2 + \right.$$

$$\left. + (12)^2 + (10)^2 + (15)^2 + (14)^2 + (19)^2 + (13)^2 + (2)^2 + (1)^2 + (4)^2 + \right.$$

$$\left. + (2)^2 + (5)^2 + (4)^2 + (5)^2 + (3)^2 \right]$$

$$SCT = 2.083 - 1.488,375 = 594,625$$

Por último, aplicamos la Fórmula (8.18) para calcular la variabilidad residual o debida al error:

$$SCR = SCT - (SCA + SCB + SCAB)$$

$$SCR = 594,625 - (445,75 + 69,46 + 37,916) = 41,5$$

Procedimiento 2: Desarrollo mediante vectores

Omitiendo el cálculo del *vector Y* que, como es sabido, es el *vector columna* de todas las *puntuaciones directas*, comenzaremos calculando los valores correspondientes a los vectores asociados a los efectos principales de los factores *A* y *B*.

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada una de las categorías del factor *A* para estimar, a partir de tales medias, los parámetros asociados al efecto del factor *A*, α_j .

$$\mu = \frac{1}{a \cdot b \cdot c} \sum_j \sum_k \sum_i Y_{ijk} \quad (8.21)$$

VECTOR B

A fin de obtener la variabilidad correspondiente al factor B , estimaremos los parámetros asociados a dicho factor, β_k . Para ello, comenzaremos calculando los promedios de las categorías del factor B .

$$\beta_k = \mu_{.k} - \mu \quad (8.24)$$

$$\mu_{.k} = \frac{1}{an} \sum_j \sum_i Y_{ijk}$$

Respectivamente:

$$\mu_{.1} = \left(\frac{1}{3 \cdot 2} \right) [4 + 3 + 14 + 12 + 2 + 1] = \frac{36}{6} = 6$$

$$\mu_{.2} = \left(\frac{1}{3 \cdot 2} \right) [7 + 6 + 12 + 10 + 4 + 2] = \frac{41}{6} = 6,833$$

$$\mu_{.3} = \left(\frac{1}{3 \cdot 2} \right) [8 + 3 + 15 + 14 + 5 + 4] = \frac{49}{6} = 8,166$$

$$\mu_{.4} = \left(\frac{1}{3 \cdot 2} \right) [12 + 11 + 19 + 13 + 5 + 3] = \frac{63}{6} = 10,5$$

Por tanto,

$$\beta_1 = \mu_{.1} - \mu = 6 - 7,875 = -1,875$$

$$\beta_2 = \mu_{.2} - \mu = 6,833 - 7,875 = -1,042$$

$$\beta_3 = \mu_{.3} - \mu = 8,166 - 7,875 = 0,291$$

$$\beta_4 = \mu_{.4} - \mu = 10,5 - 7,875 = 2,625$$

El *vector B* adopta los siguientes valores:

$$B = \{\beta\} = \begin{bmatrix} -1,875 \\ -1,875 \\ -1,042 \\ -1,042 \\ 0,291 \\ 0,291 \\ 2,625 \\ 2,625 \\ -1,875 \\ -1,875 \\ -1,042 \\ -1,042 \\ 0,291 \\ 0,291 \\ 2,625 \\ 2,625 \\ -1,875 \\ -1,875 \\ -1,042 \\ -1,042 \\ 0,291 \\ 0,291 \\ 2,625 \\ 2,625 \end{bmatrix} \quad (8.25)$$

Tras el cálculo de los vectores correspondientes a los factores A y B , procedemos a obtener el vector asociado a la interacción entre ambos factores.

VECTOR $\{\alpha\beta\}$

La variabilidad asociada al vector $\{\alpha\beta\}$ es la correspondiente al efecto de interacción entre el factor de tratamiento y el factor de bloqueo. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\alpha\beta)_{jk}$.

$$(\alpha\beta)_{jk} = \mu_{jk} - (\mu + \alpha_j + \beta_k) \quad (8.26)$$

$$\mu_{jk} = \frac{1}{n} \sum_i Y_{ijk}$$

Promedios correspondientes a las combinaciones entre los niveles de los factores A y B :

$$\mu_{11} = \left(\frac{1}{2}\right)[4 + 3] = 3,5$$

$$\mu_{12} = \left(\frac{1}{2}\right)[7 + 6] = 6,5$$

$$\mu_{13} = \left(\frac{1}{2}\right)[8 + 3] = 5,5$$

$$\mu_{14} = \left(\frac{1}{2}\right)[12 + 11] = 11,5$$

$$\mu_{21} = \left(\frac{1}{2}\right)[14 + 12] = 13$$

$$\mu_{22} = \left(\frac{1}{2}\right)[12 + 10] = 11$$

$$\mu_{23} = \left(\frac{1}{2}\right)[15 + 14] = 14,5$$

$$\mu_{24} = \left(\frac{1}{2}\right)[19 + 13] = 16$$

$$\mu_{31} = \left(\frac{1}{2}\right)[2 + 1] = 1,5$$

$$\mu_{32} = \left(\frac{1}{2}\right)[4 + 2] = 3$$

$$\mu_{33} = \left(\frac{1}{2}\right)[5 + 4] = 4,5$$

$$\mu_{34} = \left(\frac{1}{2}\right)[5 + 3] = 4$$

Por tanto:

$$(\alpha\beta)_{11} = \mu_{11} - (\mu + \alpha_1 + \beta_1) = 3,5 - (7,875 + (-1,125) + (-1,875)) = -1,375$$

$$(\alpha\beta)_{12} = \mu_{12} - (\mu + \alpha_1 + \beta_2) = 6,5 - (7,875 + (-1,125) + (-1,042)) = 0,792$$

$$(\alpha\beta)_{13} = \mu_{13} - (\mu + \alpha_1 + \beta_3) = 5,5 - (7,875 + (-1,125) + 0,291) = -1,541$$

$$(\alpha\beta)_{14} = \mu_{14} - (\mu + \alpha_1 + \beta_4) = 11,5 - (7,875 + (-1,125) + 2,625) = 2,125$$

$$(\alpha\beta)_{21} = \mu_{21} - (\mu + \alpha_2 + \beta_1) = 13 - (7,875 + 5,75 + (-1,875)) = 1,25$$

$$(\alpha\beta)_{22} = \mu_{22} - (\mu + \alpha_2 + \beta_2) = 11 - (7,875 + 5,75 + (-1,042)) = -1,583$$

$$(\alpha\beta)_{23} = \mu_{23} - (\mu + \alpha_2 + \beta_3) = 14,5 - (7,875 + 5,75 + 0,291) = 0,584$$

$$(\alpha\beta)_{24} = \mu_{24} - (\mu + \alpha_2 + \beta_4) = 16 - (7,875 + 5,75 + 2,625) = -0,25$$

$$(\alpha\beta)_{31} = \mu_{31} - (\mu + \alpha_3 + \beta_1) = 1,5 - (7,875 + (-4,625) + (-1,875)) = 0,125$$

$$(\alpha\beta)_{32} = \mu_{32} - (\mu + \alpha_3 + \beta_2) = 3 - (7,875 + (-4,625) + (-1,042)) = 0,792$$

$$(\alpha\beta)_{33} = \mu_{33} - (\mu + \alpha_3 + \beta_3) = 4,5 - (7,875 + (-4,625) + 0,291) = 0,959$$

$$(\alpha\beta)_{34} = \mu_{34} - (\mu + \alpha_3 + \beta_4) = 4 - (7,875 + (-4,625) + 2,625) = -1,875$$

El vector $\{\alpha\beta\}$ adopta los siguientes valores:

$$\{\alpha\beta\} = \begin{bmatrix} -1,375 \\ -1,375 \\ 0,792 \\ 0,792 \\ -1,541 \\ -1,541 \\ 2,125 \\ 2,125 \\ 1,25 \\ 1,25 \\ -1,583 \\ -1,583 \\ 0,584 \\ 0,584 \\ -0,25 \\ -0,25 \\ 0,125 \\ 0,125 \\ 0,792 \\ 0,792 \\ 0,959 \\ 0,959 \\ -1,875 \\ -1,875 \end{bmatrix} \quad (8.27)$$

VECTOR E, CÁLCULO DE LOS RESIDUALES O ERRORES

Despejando el término de error de la Fórmula (8.13) obtenemos:

$$\varepsilon_{ijk} = Y_{ijk} - (\mu + \alpha_j + \beta_k + (\alpha\beta)_{jk}) \quad (8.28)$$

A partir de la expresión (8.28):

$$\varepsilon_{i11} = Y_{i11} - (\mu + \alpha_1 + \beta_1 + (\alpha\beta)_{11})$$

$$\varepsilon_{i11} = Y_{i11} - (7,875 + (-1,125) + (-1,875) + (-1,375)) = Y_{i11} - 3,5$$

$$\varepsilon_{111} = Y_{111} - 3,5 = 4 - 3,5 = 0,5$$

$$\varepsilon_{211} = Y_{211} - 3,5 = 3 - 3,5 = -0,5$$

$$\varepsilon_{i12} = Y_{i12} - (\mu + \alpha_1 + \beta_2 + (\alpha\beta)_{12})$$

$$\varepsilon_{i12} = Y_{i12} - (7,875 + (-1,125) + (-1,042) + 0,792) = Y_{i12} - 6,5$$

$$\varepsilon_{112} = Y_{112} - 6,5 = 7 - 6,5 = 0,5$$

$$\varepsilon_{212} = Y_{212} - 6,5 = 6 - 6,5 = -0,5$$

$$\varepsilon_{i13} = Y_{i13} - (\mu + \alpha_1 + \beta_3 + (\alpha\beta)_{13})$$

$$\varepsilon_{i13} = Y_{i13} - (7,875 + (-1,125) + 0,291 + (-1,541)) = Y_{i13} - 5,5$$

$$\varepsilon_{113} = Y_{113} - 5,5 = 8 - 5,5 = 2,5$$

$$\varepsilon_{213} = Y_{213} - 5,5 = 3 - 5,5 = -2,5$$

$$\varepsilon_{i14} = Y_{i14} - (\mu + \alpha_1 + \beta_4 + (\alpha\beta)_{14})$$

$$\varepsilon_{i14} = Y_{i14} - (7,875 + (-1,125) + 2,625 + 2,125) = Y_{i14} - 11,5$$

$$\varepsilon_{114} = Y_{114} - 11,5 = 12 - 11,5 = 0,5$$

$$\varepsilon_{214} = Y_{214} - 11,5 = 11 - 11,5 = -0,5$$

$$\varepsilon_{i21} = Y_{i21} - (\mu + \alpha_2 + \beta_1 + (\alpha\beta)_{21})$$

$$\varepsilon_{i21} = Y_{i21} - (7,875 + 5,75 + (-1,875) + 1,25) = Y_{i21} - 13$$

$$\varepsilon_{121} = Y_{121} - 13 = 14 - 13 = 1$$

$$\varepsilon_{221} = Y_{221} - 13 = 12 - 13 = -1$$

$$\varepsilon_{i22} = Y_{i22} - (\mu + \alpha_2 + \beta_2 + (\alpha\beta)_{22})$$

$$\varepsilon_{i22} = Y_{i22} - (7,875 + 5,75 + (-1,042) + (-1,538)) = Y_{i22} - 11$$

$$\varepsilon_{122} = Y_{122} - 11 = 12 - 11 = 1$$

$$\varepsilon_{222} = Y_{222} - 11 = 10 - 11 = -1$$

$$\varepsilon_{i23} = Y_{i23} - (\mu + \alpha_2 + \beta_3 + (\alpha\beta)_{23})$$

$$\varepsilon_{i23} = Y_{i23} - (7,875 + 5,75 + 0,291 + 0,584) = Y_{i23} - 14,5$$

$$\varepsilon_{123} = Y_{123} - 14,5 = 15 - 14,5 = 0,5$$

$$\varepsilon_{223} = Y_{223} - 14,5 = 14 - 14,5 = -0,5$$

$$\varepsilon_{i24} = Y_{i24} - (\mu + \alpha_2 + \beta_4 + (\alpha\beta)_{24})$$

$$\varepsilon_{i24} = Y_{i24} - (7,875 + 5,75 + 2,625 + (-0,25)) = Y_{i24} - 16$$

$$\varepsilon_{124} = Y_{124} - 16 = 19 - 16 = 3$$

$$\varepsilon_{224} = Y_{224} - 16 = 13 - 16 = -3$$

$$\varepsilon_{i31} = Y_{i31} - (\mu + \alpha_3 + \beta_1 + (\alpha\beta)_{31})$$

$$\varepsilon_{i31} = Y_{i31} - (7,875 + (-4,625) + (-1,875) + 0,125) = Y_{i31} - 1,5$$

$$\varepsilon_{131} = Y_{131} - 1,5 = 2 - 1,5 = 0,5$$

$$\varepsilon_{231} = Y_{231} - 1,5 = 1 - 1,5 = -0,5$$

$$\varepsilon_{i32} = Y_{i32} - (\mu + \alpha_3 + \beta_2 + (\alpha\beta)_{32})$$

$$\varepsilon_{i32} = Y_{i32} - (7,875 + (-4,625) + (-1,042) + 0,792) = Y_{i32} - 3$$

$$\varepsilon_{132} = Y_{132} - 3 = 4 - 3 = 1$$

$$\varepsilon_{232} = Y_{232} - 3 = 2 - 3 = -1$$

$$\varepsilon_{i33} = Y_{i33} - (\mu + \alpha_3 + \beta_3 + (\alpha\beta)_{33})$$

$$\varepsilon_{i33} = Y_{i33} - (7,875 + (-4,625) + 0,291 + 0,959) = Y_{i33} - 4,5$$

$$\varepsilon_{133} = Y_{133} - 4,5 = 5 - 4,5 = 0,5$$

$$\varepsilon_{233} = Y_{233} - 4,5 = 4 - 4,5 = -0,5$$

$$\varepsilon_{i34} = Y_{i34} - (\mu + \alpha_3 + \beta_4 + (\alpha\beta)_{34})$$

$$\varepsilon_{i34} = Y_{i34} - (7,875 + (-4,625) + 2,625 + (-1,875)) = Y_{i34} - 4$$

$$\varepsilon_{134} = Y_{134} - 4 = 5 - 4 = 1$$

$$\varepsilon_{234} = Y_{234} - 4 = 3 - 4 = -1$$

Por tanto, el vector E adopta los siguientes valores:

$$E = \{\varepsilon\} = \begin{bmatrix} 0,5 \\ -0,5 \\ 0,5 \\ -0,5 \\ 2,5 \\ -2,5 \\ 0,5 \\ -0,5 \\ 1 \\ -1 \\ 1 \\ -1 \\ 0,5 \\ -0,5 \\ 3 \\ -3 \\ 0,5 \\ -0,5 \\ 1 \\ -1 \\ 0,5 \\ -0,5 \\ 1 \\ -1 \end{bmatrix} \quad (8.29)$$

Llegados a este punto, podemos calcular la *SCA*, la *SCB*, la *SCAB*, la *SCR* y la *SCT* aplicando las Fórmulas (8.30), (8.31), (8.32), (8.33) y (8.34), respectivamente, o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos:

VARIABILIDAD CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = n \cdot b \sum_j \alpha_j^2 \quad (8.30)$$

$$SCA = 2 \cdot 4[(-1,125)^2 + (5,75)^2 + (-4,625)^2] = 445,75$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR B (SCB):

$$SCB = n \cdot a \sum_k \beta_k^2 \quad (8.31)$$

$$SCB = 2 \cdot 3[(-1,875)^2 + (-1,042)^2 + (0,291)^2 + (2,625)^2] = 69,46$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y B (SCAB):

$$SCAB = n \sum_j \sum_k (\alpha\beta)_{jk}^2 \quad (8.32)$$

$$SCAB = 2 \left[\begin{array}{l} (-1,375)^2 + (0,792)^2 + (-1,541)^2 + (2,125)^2 + (1,250)^2 + \\ + (-1,583)^2 + (0,584)^2 + (-0,250)^2 + (0,125)^2 + \\ + (0,792)^2 + (0,959)^2 + (-1,875)^2 \end{array} \right] = 37,916$$

VARIABILIDAD RESIDUAL (SCR):

$$SCR = \sum_i \sum_j \sum_k \varepsilon_{ijk}^2 \quad (8.33)$$

$$SCR = \left[\begin{array}{l} (0,5)^2 + (-0,5)^2 + (0,5)^2 + (-0,5)^2 + (2,5)^2 + (-2,5)^2 + \\ + (0,5)^2 + (-0,5)^2 + (1)^2 + (-1)^2 + (1)^2 + (-1)^2 + (0,5)^2 + \\ + (-0,5)^2 + (3)^2 + (-3)^2 + (0,5)^2 + (-0,5)^2 + (1)^2 + \\ + (-1)^2 + (0,5)^2 + (-0,5)^2 + (1)^2 + (-1)^2 \end{array} \right] = 41,5$$

VARIABILIDAD TOTAL (SCT):

$$SCT = SCA + SCB + SCAB + SCR \quad (8.34)$$

$$SCT = 445,75 + 69,46 + 37,916 + 41,5 = 594,626$$

Multiplicando cada vector por su traspuesto obtenemos los mismos resultados:

$$SCA = \{\alpha\}^T \{\alpha\} = 445,75 \quad (8.35) \quad SCAB = \{\alpha\beta\}^T \{\alpha\beta\} = 37,916 \quad (8.37)$$

$$SCB = \{\beta\}^T \{\beta\} = 69,46 \quad (8.36) \quad SCR = \{\varepsilon\}^T \{\varepsilon\} = 41,5 \quad (8.38)$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 8.8 Análisis factorial de la varianza para el diseño de bloques aleatorios con un factor de tratamiento y un factor de bloqueo y con más de un sujeto por casilla: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Entre tratamientos (factor A)	$SCA = 445,75$	$a - 1 = 2$	$MCA = 222,875$	$F_A = 64,41$
Entre bloques (factor B)	$SCB = 69,46$	$b - 1 = 3$	$MCB = 23,15$	$F_B = 6,69$
Interacción $A \times B$	$SCAB = 37,916$	$(a - 1)(b - 1) = 6$	$MCAB = 6,32$	$F_{AB} = 1,83$
Error	$SCR = 41,5$	$ab(n - 1) = 12$	$MCR = 3,46$	
Total	$SCT = 594,626$	$abn - 1 = 23$		

Tras la obtención de las F observadas, debemos determinar si la variabilidad explicada por los factores y por su interacción es o no significativa. Para ello, recurrimos a las *tablas de los valores críticos de la distribución F*. Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos (véase la Tabla 8.9).

TABLA 8.9 Comparaciones entre las F observadas y las F teóricas con un nivel de confianza del 95 %

Fuentes de variación	F crítica _(0,95; gl1/gl2)	F observada	Diferencia
Factor A	$F_{0,95;2/12} = 3,89$	$F_A = 64,41$	$64,41 > 3,88$
Factor B	$F_{0,95;3/12} = 3,49$	$F_B = 6,69$	$6,69 > 3,49$
Interacción $A \times B$	$F_{0,95;6/12} = 3$	$F_{AB} = 1,83$	$1,83 < 3$

Como puede apreciarse en la Tabla 8.9, rechazamos la hipótesis nula para el efecto de la *metodología de trabajo empleada en el aula* (factor A). A su vez, también cabe concluir que el factor de bloqueo o la *motivación con respecto al aprendizaje* aporta al modelo una fuente de variación, cuya influencia sobre la variable criterio resulta estadísticamente significativa. Por otra parte, no existe efecto de interacción entre la variable de tratamiento y la de bloqueo. Estos dos últimos resultados permiten afirmar que en el presente experimento la técnica de bloqueo ha sido correctamente aplicada, ya que coinciden con los dos requisitos fundamentales para la utilización de dicha estrategia, a saber, la ausencia de interacción entre la variable independiente y la variable bloqueada, y la existencia de relación entre la variable bloqueada y la variable dependiente. El hecho de que se satisfaga esta segunda condición permite que el bloqueo cumpla con su finalidad básica, la cual consiste en aumentar la sensibilidad del diseño o, dicho de otro modo, en aumentar la precisión en la estimación de los efectos experimentales.

• Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

El análisis de la varianza para un diseño de bloques se realiza mediante el mismo procedimiento aplicado al diseño factorial $A \times B$, expuesto en el Epígrafe 7.3.2.4.

La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo sería:

```
UNIANOVA
  rendimie BY metodol motivaci
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /EMMEANS = TABLES(metodol)
  /EMMEANS = TABLES(motivaci)
  /PRINT = DESCRIPTIVE ETASQ OPOWER
  /CRITERIA = ALPHA(.05)
  /DESIGN = metodol motivaci metodol*motivaci.
```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Metodología de trabajo	1,00	Audio-visual	8
	2,00	Naturaleza	8
	3,00	Tradicional	8
Nivel de motivación	1,00	Bajo	6
	2,00	Medio-bajo	6
	3,00	Medio-alto	6
	4,00	Alto	6

Estadísticos descriptivos

Variable dependiente: Rendimiento

Metodología de trabajo	Nivel de motivación	Media	Desv. típ.	N
Audio-visual	Bajo	3,5000	0,7071	2
	Medio-bajo	6,5000	0,7071	2
	Medio-alto	5,5000	3,5355	2
	Alto	11,5000	0,7071	2
	Total	6,7500	3,4538	8
Naturaleza	Bajo	13,0000	1,4142	2
	Medio-bajo	11,0000	1,4142	2
	Medio-alto	14,5000	0,7071	2
	Alto	16,0000	4,2426	2
	Total	13,6250	2,6693	8
Tradicional	Bajo	1,5000	0,7071	2
	Medio-bajo	3,0000	1,4142	2
	Medio-alto	4,5000	0,7071	2
	Alto	4,0000	1,4142	2
	Total	3,2500	1,4880	8
Total	Bajo	6,0000	5,5498	6
	Medio-bajo	6,8333	3,7103	6
	Medio-alto	8,1667	5,1929	6
	Alto	10,5000	5,7879	6
	Total	7,8750	5,0846	24

Pruebas de los efectos intersujetos**Variable dependiente: Rendimiento**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	553,125 ^b	11	50,284	14,54	0,000	0,930	159,940	1,000
Intercept	1.488,375	1	1.488,375	430,4	0,000	0,973	430,373	1,000
METODOL.	445,750	2	222,875	64,45	0,000	0,915	128,892	1,000
MOTIVACI.	69,458	3	23,153	6,695	0,007	0,626	20,084	0,904
METODOL. * MOTIVACI.	37,917	6	6,319	1,827	0,176	0,477	10,964	0,459
Error	41,500	12	3,458					
Total	2.083,000	24						
Total corregido	594,625	23						

^a Calculado con alfa = 0,05.^b R cuadrado = 0,930 (R cuadrado corregido = 0,866).**Medias marginales estimadas****1. Metodología de trabajo****Variable dependiente: Rendimiento**

Metodología de trabajo	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Audio-visual	6,750	0,657	5,317	8,183
Naturaleza	13,625	0,657	12,192	15,058
Tradicional	3,250	0,657	1,817	4,683

2. Nivel de motivación**Variable dependiente: Rendimiento**

Metodología de trabajo	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Bajo	6,000	0,759	4,346	7,654
Medio-bajo	6,833	0,759	5,179	8,487
Medio-alto	8,167	0,759	6,513	9,821
Alto	10,500	0,759	8,846	12,154

8.2. DISEÑOS CON DOS FACTORES DE BLOQUEO:
DISEÑOS DE CUADRADO LATINO

8.2.1. Características generales del diseño de cuadrado latino

En esta categoría se incluyen los diseños que utilizan dos sistemas de bloques simultáneamente. En tales diseños los niveles de un factor experimental se administran en cada una de las categorías de dos factores de bloqueo, con el objetivo de controlar dos variables extrañas. Sin embargo, no se generan tantas condiciones experimentales como combinaciones posibles entre los factores, ya que ello aumentaría considerablemente el coste del experimento. Así, con la finalidad de aprovechar las ventajas de la técnica de bloqueo, sin incrementar excesivamente la complejidad de la investigación, se han elaborado diseños de matriz cuadrada en los que las dimensiones de los bloques coinciden con el número de tratamientos experimentales. Cabe señalar que, dentro de este tipo de estructuras de investigación, se incluyen, además del diseño de cuadrado latino, otras modalidades de diseño con más de una dimensión de bloqueo. En concreto, los diseños de este tipo, que aplican la *técnica de doble bloqueo*, se conocen como *diseños de cuadrado latino*. Aquellos que utilizan tres y cuatro sistemas de bloques se denominan *diseños de cuadrado grecolatino* y *diseños de cuadrado hipergrecolatino*, respectivamente. Dado que la lógica que subyace a todos los *diseños con más de una dimensión de bloqueo* es la misma, en el presente epígrafe abordaremos la estructura más simple, a saber, el *diseño de cuadrado latino intersujetos*.

Un **diseño de cuadrado latino** es un diseño de bloques con un factor experimental y dos variables de bloqueo. En este tipo de diseño, la muestra de sujetos se estratifica en función de dos variables de clasificación y, posteriormente, se aplican los distintos tratamientos dentro de cada bloque. En consecuencia, se combinan tres dimensiones de variación: la debida a los tratamientos y las dos dimensiones de variación debidas a las dos variables de bloqueo, las cuales coinciden con las filas y con las columnas de una matriz cuadrada o de doble entrada. Cada una de estas dimensiones actúa a k niveles. El procedimiento de doble bloqueo permite obtener muestras de sujetos muy homogéneas, con lo que se reduce en gran medida la varianza de error asociada a las diferencias individuales y se incrementa la precisión en la estimación de los efectos de la variable de tratamiento.

El diseño de cuadrado latino se puede definir como una modalidad de diseño factorial fraccionado que tiene la misma cantidad de factores y de niveles, y en el que las interacciones se consideran nulas (Arnau, 1986). Dado que los factores no se cruzan totalmente, este tipo de diseño sólo permite estimar los efectos principales y el término de error. Cuando se utiliza, por ejemplo, un diseño factorial completo que consta de un factor experimental (A) con tres niveles (a_1, a_2, a_3) y de dos factores de clasificación (B y C) con otras tres categorías cada uno (b_1, b_2, b_3 y c_1, c_2, c_3 , respectivamente), el total de condiciones experimentales es de 27 ($3 \times 3 \times 3 = 3^3 = 27$). En el diseño de cuadrado latino las condiciones de tratamiento se reducen a 9, lo que constituye una réplica de $1/3$ del total de combinaciones que se derivan del correspondiente diseño factorial.

El diseño de cuadrado latino 3×3 se puede representar mediante la siguiente matriz de doble entrada:

	C_1	C_2	C_3
B_1	A_1	A_2	A_3
B_2	A_2	A_3	A_1
B_3	A_3	A_1	A_2

Figura 8.1 Representación del diseño de cuadrado latino 3×3 mediante una matriz de doble entrada.

Cabe señalar que esta matriz cuadrada constituye tan sólo una de las posibles configuraciones que puede presentar un diseño de cuadrado latino. De hecho, para cada tamaño de la matriz cuadrada o del diseño, existen disposiciones que se consideran *prototípicas* o *estándar*. El *cuadrado latino estándar* se define como aquella disposición en la que, tanto la primera fila como la primera columna siguen el orden o la secuencia numérica natural. A partir de cada una de las disposiciones estándar se puede obtener una determinada cantidad de formatos diferentes que se derivan de las posibles combinaciones entre filas y columnas. La disposición cuadrada 3×3 , por ejemplo, sólo posee un formato estándar que coincide con el formato que hemos presentado como ejemplo de cuadrado latino. A partir de esta forma estándar, pueden obtenerse $(3!) \times (2!) - 1 = 11$ diferentes cuadrados latinos 3×3 , que corresponden a las posibles combinaciones entre las distintas filas y columnas de la matriz. Si a estas once configuraciones se les añade el formato estándar, obtenemos un total de 12 cuadrados latinos 3×3 . Como cabe deducir de la fórmula que se emplea para calcular la cantidad de formatos que se derivan de la disposición prototípica, el número de posibles formatos aumenta rápidamente a medida que la dimensión del cuadrado latino toma más valores. Así, por ejemplo, de una matriz de 4×4 se pueden obtener hasta 576 cuadrados latinos, a partir de los cuatro formatos estándar correspondientes a dicha matriz. Los textos de Fisher y Yates (1953) y de Cochran y Cox (1957) proporcionan una serie de tablas en las que se presentan las principales formas estándar de los cuadrados latinos. Es muy importante señalar que, independientemente de la disposición que adopte el cuadrado latino, cada valor de la variable de tratamiento debe aparecer una sola vez en cada fila y en cada columna de la matriz.

Además de la condición que se acaba de citar, el diseño de cuadrado latino debe cumplir otra serie de requisitos que garantizan su correcta aplicación:

- El diseño debe ser *equilibrado*. En caso contrario, el efecto de las variables de bloqueo no permanece constante a lo largo de todos los tratamientos.
- Las interacciones no deben ejercer influencia significativa sobre la variable de respuesta ya que, como se ha señalado anteriormente, el diseño de cuadrado latino no permite examinar tales interacciones. Si no se cumple este requisito resulta más adecuado plantear un diseño factorial completo.
- Las variables de bloqueo deben guardar una estrecha relación con la variable dependiente. Esta condición es realmente importante, ya que la efectividad del diseño de cuadrado latino depende del grado de relación existente entre las variables de bloqueo y la variable de respuesta.

8.2.2. Análisis factorial de la varianza para el diseño de cuadrado latino

8.2.2.1. Modelo general de análisis

Al igual que en el diseño de bloques aleatorios, el modelo analítico que se utiliza habitualmente para llevar a cabo la **prueba de la hipótesis** en el diseño de cuadrado latino es el *análisis factorial de la varianza*.

Suponiendo una **estructura básica de tres factores y de un solo sujeto por casilla**, el modelo matemático estructural que subyace al análisis de la varianza, bajo el supuesto de la hipótesis alternativa, responde a la siguiente expresión:

$$y_{ijkm} = \mu + \alpha_j + \beta_k + \gamma_m + \varepsilon_{ijkm} \quad (8.39)$$

donde:

y_{ijkm} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable A o variable de tratamiento, la k -ésima fila de la variable B o variable de bloqueo y la m -ésima columna de la variable C o variable de bloqueo.

μ = Media general de la variable dependiente.

α_j = Efecto principal asociado al j -ésimo nivel de la variable A o variable de tratamiento.

β_k = Efecto principal asociado a la k -ésima fila de la variable B o variable de bloqueo.

γ_m = Efecto principal asociado a la m -ésima columna de la variable C o variable de bloqueo.

ε_{ijkm} = Término de error o componente aleatorio del modelo.

Se asume que:

a) $\varepsilon_{ijkm} \simeq NID(0, \sigma^2)$.

b) Las interacciones $(\alpha\beta)_{jk}$, $(\alpha\gamma)_{jm}$, $(\beta\gamma)_{km}$ y $(\alpha\beta\gamma)_{jkm}$ son nulas.

En este caso, la variación total se descompone en cuatro fuentes de variación: la debida a la variable de tratamiento, las variaciones asociadas a las dos variables de bloqueo y la debida al componente residual del modelo. Las tres primeras configuran la *variación inter-grupos* y la cuarta corresponde a la *variación intragrupo*. A partir de estas variaciones o sumas de cuadrados, pueden obtenerse las varianzas y las razones F que permiten estimar los efectos de los respectivos componentes de la ecuación estructural del diseño.

En relación con el análisis de la varianza en el diseño de cuadrado latino, es necesario hacer algunas **consideraciones importantes**. La primera de ellas se refiere a los grados de libertad asociados al componente de error. De hecho, uno de los mayores inconvenientes que podemos encontrar, al utilizar matrices cuadradas de dimensiones pequeñas, radica en la reducida cantidad de grados de libertad que se asocian al término de error. Ello constituye un serio problema porque, como afirma Edwards (1985), una estimación del error basada en menos de 30 grados de libertad no es fiable. Ante este problema pueden adoptarse dos soluciones. Por un lado, cabe incrementar el número de grados de libertad del error llevando a cabo varias réplicas diferentes del mismo diseño. Por otro lado, el incremento en los grados de libertad también puede lograrse aumentando el número de sujetos en cada combinación de tratamientos. Aunque la primera estrategia resulta muy deseable, es difícil de aplicar. Por ello, Arnau (1986) aconseja ampliar el tamaño de la muestra hasta obtener una cantidad de grados de libertad adecuada para poder estimar con precisión los efectos de los tratamientos experimentales.

La segunda consideración se refiere al tipo de matriz utilizada para llevar a cabo el experimento. A este respecto, Pascual (1995b) señala que, si se emplea una matriz estándar y de naturaleza cíclica, nos podemos encontrar con un patrón sistemático de orden que convierte el diseño en una estructura vulnerable a los efectos que se conocen como *efectos carry-over*. Como veremos al abordar los diseños de medidas repetidas, dichos efectos se producen cuando el efecto de un tratamiento no ha desaparecido en el momento en el que se introduce el siguiente tratamiento y, en consecuencia, ejerce influencia sobre las respuestas dadas por los sujetos en este último tratamiento. En tal caso, resulta conveniente utilizar diferentes matrices de cuadrado latino con diferentes grupos de sujetos, elevando el número de unidades experimentales disponibles a un múltiplo de los a tratamientos. Ya se ha señalado anteriormente que la realización de réplicas completas del experimento no es tarea fácil.

No obstante, constituye un procedimiento muy útil, tanto para evitar los efectos *carry-over*, como para incrementar los grados de libertad del término de error.

8.2.2.2. Ejemplo práctico

Supongamos que en el ámbito de la psicología de las organizaciones, realizamos una investigación para examinar la influencia que ejerce el *nivel de responsabilidad asumido en la empresa* (factor A) sobre la *satisfacción laboral*. No obstante, deseamos bloquear, simultáneamente, el efecto de la *edad del trabajador* (factor B) y del *turno horario* (factor C). A tal fin, se escoge una muestra compuesta por 9 sujetos y se establece, aleatoriamente, que tres de ellos asuman, durante un determinado período temporal, un *nivel de responsabilidad alto* (a_1), otros tres un *nivel de responsabilidad medio* (a_2) y los tres restantes un *nivel de responsabilidad bajo* (a_3). Como ya se ha señalado, tales sujetos pertenecen a tres grupos de edad diferentes, a saber, (b_1) 20-35 años, (b_2) 36-50 años y (b_3) más de 51 años. A su vez, trabajan en uno de los tres siguientes turnos horarios: (c_1) *mañana*, (c_2) *tarde* y (c_3) *noche*. En la Tabla 8.10 pueden observarse las puntuaciones obtenidas por los sujetos en un cuestionario destinado a medir la satisfacción laboral tras permanecer dos meses en la empresa.

TABLA 8.10 Matriz de datos del experimento

		C (Turno horario)			Sumatorios y medias marginales	A (Nivel de responsabilidad)
		c_1 (Mañana)	c_2 (Tarde)	c_3 (Noche)		
B (Edad)	b_1 (20-35)	a_1 (Alto) 20	a_2 (Medio) 16	a_3 (Bajo) 12	$\Sigma Y_{.1} = 48$ $\bar{Y}_{.1} = 16$	$\Sigma Y_{1..} = 52$ $\bar{Y}_{1..} = 17,33$
	b_2 (36-50)	a_2 (Medio) 12	a_3 (Bajo) 10	a_1 (Alto) 14	$\Sigma Y_{.2} = 36$ $\bar{Y}_{.2} = 12$	$\Sigma Y_{2..} = 38$ $\bar{Y}_{2..} = 12,66$
	b_3 (Más de 51)	a_3 (Bajo) 8	a_1 (Alto) 18	a_2 (Medio) 10	$\Sigma Y_{.3} = 36$ $\bar{Y}_{.3} = 12$	$\Sigma Y_{3..} = 30$ $\bar{Y}_{3..} = 10$
Sumatorios y medias marginales		$\Sigma Y_{..1} = 40$ $\bar{Y}_{..1} = 13,33$	$\Sigma Y_{..2} = 44$ $\bar{Y}_{..2} = 14,66$	$\Sigma Y_{..3} = 36$ $\bar{Y}_{..3} = 12$	$\Sigma Y_{...} = 120$ $\bar{Y}_{...} = 13,33$	

Dado que el objetivo del presente subapartado consiste, fundamentalmente, en ilustrar la técnica y la modelización del diseño de cuadrado latino, utilizaremos únicamente el primero de los procedimientos que hemos empleado para el cálculo de las sumas de cuadrados en los diseños abordados con anterioridad y, a partir de tales sumas cuadráticas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la ecuación estructural del diseño (véase la Fórmula 8.39). Cabe señalar que el desarrollo del ANOVA mediante vectores no plantea ninguna dificultad añadida y se realiza exactamente igual que en el caso de los diseños que hemos expuesto en los epígrafes precedentes. Tras estas aclaraciones, procedemos a desarrollar el análisis factorial de la varianza para el diseño de cuadrado latino

con un solo sujeto por casilla bajo el modelo de aditividad, tomando como referencia la matriz de datos de la Tabla 8.10.

En la Tabla 8.11 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas.

TABLA 8.11 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño de cuadrado latino con un solo sujeto por casilla bajo el modelo de aditividad

Efecto del factor de tratamiento (factor A)	$SCA = \left[\frac{1}{a} \sum_j \left(\sum_i \sum_k \sum_m Y_{ijkm} \right)^2 \right] - C \quad (8.40)$
Efecto del primer factor de bloqueo (factor B)	$SCB = \left[\frac{1}{a} \sum_k \left(\sum_i \sum_j \sum_m Y_{ijkm} \right)^2 \right] - C \quad (8.41)$
Efecto del segundo factor de bloqueo (factor C)	$SCC = \left[\frac{1}{a} \sum_m \left(\sum_i \sum_j \sum_k Y_{ijkm} \right)^2 \right] - C \quad (8.42)$
Variabilidad residual o del error	$SCR = SCT - SCA - SCB - SCC \quad (8.43)$
Variabilidad total	$SCT = \sum_i \sum_j \sum_k \sum_m Y_{ijkm}^2 - C \quad (8.44)$
C	$C = \frac{1}{a^2} \left(\sum_i \sum_j \sum_k \sum_m Y_{ijkm} \right)^2 \quad (8.45)$

Tras obtener el valor de C, calcularemos las diferentes sumas de cuadrados:

$$C = \frac{1}{3^2} (120)^2 = 1.600$$

Una vez obtenido este valor, procedemos a calcular la SCA, la SCB, la SCC, la SCT y la SCR, respectivamente:

$$SCA = \frac{1}{3} [(52)^2 + (38)^2 + (30)^2] - 1.600 = 1.682,66 - 1.600 = 82,66$$

$$SCB = \frac{1}{3} [(48)^2 + (36)^2 + (36)^2] - 1.600 = 1.632 - 1.600 = 32$$

$$SCC = \frac{1}{3} [(40)^2 + (44)^2 + (36)^2] - 1.600 = 1.610,66 - 1.600 = 10,66$$

$$SCT = [(20)^2 + (12)^2 + (8)^2 + (16)^2 + (10)^2 + (18)^2 + (12)^2 + (14)^2 + (10)^2] - 1.600 = 1.728 - 1.600 = 128$$

$$SCR = 128 - 82,66 - 32 - 10,66 = 2,68$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 8.12 Análisis factorial de la varianza para el diseño de cuadrado latino intersujetos con un solo sujeto por casilla bajo el modelo de aditividad: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Entre tratamientos (factor A)	SCA = 82,66	$a - 1 = 2$	MCA = 41,33	$F_A = 30,84$
Entre filas (factor B)	SCB = 32	$a - 1 = 2$	MCB = 16	$F_B = 11,94$
Entre columnas (factor C)	SCC = 10,66	$a - 1 = 2$	MCC = 5,33	$F_C = 3,98$
Residual o del error	SCR = 2,68	$(a - 1)(a - 2) = 2$	MCR = 1,34	
Total	SCT = 128	$a^2 - 1 = 8$		

Tras consultar las *tablas de los valores críticos de la distribución F* con un nivel de confianza del 95 % ($\alpha = 0,05$) y trabajando con una hipótesis de una cola, podemos concluir que el *nivel de responsabilidad asumido en la empresa* (factor A) ejerce una influencia estadísticamente significativa sobre la *satisfacción laboral* ($F_{0,95;2,2} = 19 < F_{\text{obs.}} = 30,84$). Por otra parte, mantenemos la hipótesis nula tanto para el efecto del factor B ($F_{0,95;2,2} = 19 > F_{\text{obs.}} = 11,94$) como para el efecto del factor C ($F_{0,95;2,2} = 19 > F_{\text{obs.}} = 3,98$).

Examinemos cuáles habrían sido los resultados del análisis estadístico en el caso de no haber bloqueado el efecto de los factores B y C, es decir, en el caso de haber utilizado un diseño multigrupos aleatorios en lugar de un diseño de cuadrado latino.

Para ello, redefinimos el término de error del ANOVA incluyendo, dentro de dicho término, el efecto de las variables bloqueadas. De esta forma, obtenemos la siguiente tabla del análisis de la varianza.

TABLA 8.13 Análisis unifactorial de la varianza para un diseño multigrupos aleatorios

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A	SCA = 82,66	2	MCA = 41,33	$F_A = 5,47$
Residual o del error	SCR = 45,34	6	MCR = 7,56	
Total	SCT = 128	8		

Como cabía esperar, teniendo en cuenta que los factores de bloqueo no ejercen influencia sobre la variable criterio, se siguen apreciando diferencias estadísticamente significativas en la *satisfacción laboral* en función del *nivel de responsabilidad asumido en la empresa* ($F_{0,95;2,6} = 5,14 < F_{\text{obs.}} = 5,47$). No obstante, es evidente que el diseño de cuadrado latino consigue reducir, considerablemente, la varianza de error. Además, la reducción de dicha varianza es mayor a medida que aumenta la relación existente entre las variables de bloqueo y la variable dependiente.

8.2.2.3. *El problema de la reducción de los grados de libertad del término de error*

Al abordar el modelo general de análisis factorial de la varianza, para el diseño de cuadrado latino, se han señalado dos inconvenientes importantes que todo investigador debe considerar si trabaja con este modelo de diseño, a saber, el referido a la reducción que se produce en el número de grados de libertad del componente de error, cuando se utilizan matrices cuadradas de dimensiones pequeñas, y la vulnerabilidad de este tipo de diseños ante los efectos *carry-over*. También se han señalado las posibles soluciones para ambos tipos de problemas. Así, mientras la realización de réplicas completas del experimento permite solventar los dos inconvenientes, el incremento en el número de sujetos por combinación de tratamientos responde al primero de ellos. Dado que la reducción de los grados de libertad del término de error es uno de los problemas más graves asociados al ANOVA para el diseño de cuadrado latino, consideraremos, brevemente, cada una de las alternativas de las que disponemos para solucionarlo.

• **Aumento de la cantidad de réplicas completas del experimento**

Tomando como referencia el ejemplo que hemos empleado para ilustrar el análisis factorial de la varianza en el diseño de cuadrado latino, la obtención de la cantidad de grados de libertad que la mayoría de los investigadores consideran necesaria para realizar una adecuada estimación del error nos obligaría a llevar a cabo 15 réplicas diferentes del mismo experimento. Los grados de libertad quedarían distribuidos tal y como se muestra en la Tabla 8.14.

TABLA 8.14 Distribución de los grados de libertad de un diseño de cuadrado latino 3×3 con un solo sujeto por casilla y 15 réplicas del experimento

Fuentes de variación	Grados de libertad
Entre tratamientos (factor A)	$r(a - 1) = 15(3 - 1) = 30$
Entre filas (factor B)	$r(a - 1) = 15(3 - 1) = 30$
Entre columnas (factor C)	$r(a - 1) = 15(3 - 1) = 30$
Entre cuadrados latinos (réplicas)	$r - 1 = 15 - 1 = 14$
Residual o del error	$r(a - 1)(a - 2) = 15(3 - 1)(3 - 2) = 30$
Total	$ra^2 - 1 = 15 \cdot 9 - 1 = 134$

• **Aumento del número de sujetos por combinación de tratamientos**

Partiendo del experimento propuesto a lo largo del presente epígrafe (Epígrafe 8.2.2), la obtención del número de grados de libertad necesario para que la estimación del error resulte fiable, nos obligaría a utilizar, como mínimo, 5 sujetos en cada combinación de tratamientos. De esta forma, los grados de libertad quedarían distribuidos tal y como se muestra en la Tabla 8.15.

TABLA 8.15 Distribución de los grados de libertad de un diseño de cuadrado latino 3×3 con cinco sujetos por casilla

Fuentes de variación	Grados de libertad
Entre tratamientos (factor A)	$a - 1 = 3 - 1 = 2$
Entre filas (factor B)	$a - 1 = 3 - 1 = 2$
Entre columnas (factor C)	$a - 1 = 3 - 1 = 2$
No aditividad (residual entre celdas)	$(a - 1)(a - 2) = (3 - 1)(3 - 2) = 2$
Entre celdas	$(a - 1)(n - 1) = (3 - 1)(5 - 1) = 8$
Error (intra celdas)	$a^2(n - 1) = 9(5 - 1) = 36$
Total	$a^2n - 1 = 44$

8.2.2.4. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la(s) variable(s) dependiente(s) y los factores, especificando en su lugar correspondiente si tales factores son fijos o aleatorios (en nuestro ejemplo, los tres factores son fijos).



- El menú **Modelo** nos permite especificar el modelo que, por defecto, es un modelo factorial completo. En nuestro caso, escogemos la opción *Personalizado*. A continuación, nos situamos en *Construir términos* y seleccionamos la opción *Efectos princip.* del menú desplegable. Seguidamente, tras escoger los tres factores de nuestro ejemplo, los colocamos en el cuadro derecho donde se indica *Modelo*, utilizando el botón destinado a tal fin:



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar «estimaciones del tamaño del efecto»* y *«potencia observada»*).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

UNIANOVA

```
satisfac BY edad turno respons
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(respons)
/EMMEANS = TABLES(edad)
/EMMEANS = TABLES(turno)
/PRINT = ETASQ OPOWER
/CRITERIA = ALPHA(.05)
/DESIGN = respons edad turno.
```


- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Edad	1,00	20-35	3
	2,00	36-50	3
	3,00	Más de 51	3
Turno horario	1,00	Mañana	3
	2,00	Tarde	3
	3,00	Noche	3
Nivel de responsabilidad	1,00	Alto	3
	2,00	Medio	3
	3,00	Bajo	3

Pruebas de los efectos intersujetos

Variable dependiente: Satisfacción laboral

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	125,333 ^b	6	20,889	15,667	0,061	0,979	94,000	0,572
Intersección	1.600,000	1	1.600,000	1.200,0	0,001	0,998	1.200,000	1,000
RESPONS.	82,667	2	41,333	31,000	0,031	0,969	62,000	0,798
EDAD	32,000	2	16,000	12,000	0,007	0,923	24,000	0,479
TURN0	10,667	2	5,333	4,000	0,200	0,800	8,000	0,222
Error	2,667	2	1,333					
Total	1.728,000	9						
Total corregido	128,000	8						

^a Calculado con alfa = 0,05.

^b R cuadrado = 0,979 (R cuadrado corregido = 0,917).

Medias marginales estimadas

1. Nivel de responsabilidad

Variable dependiente: Satisfacción laboral

Nivel de responsabilidad	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Alto	17,333	0,667	14,465	20,202
Medio	12,667	0,667	9,798	15,535
Bajo	10,000	0,667	7,132	12,868

2. Edad

Variable dependiente: Satisfacción laboral

Edad	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
20-35	16,000	0,667	13,132	18,868
36-50	12,000	0,667	9,132	14,868
Más de 51	12,000	0,667	9,132	14,868

3. Turno horario

Variable dependiente: Satisfacción laboral

Turno horario	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Mañana	13,333	0,667	10,465	16,202
Tarde	14,667	0,667	11,798	17,535
Noche	12,000	0,667	9,132	14,868

8.3. DISEÑOS JERÁRQUICOS

8.3.1. Características generales del diseño jerárquico

El *diseño jerárquico*, conocido también como *diseño con variables anidadas* o *diseño de sujetos intragrupos intratratamientos*, es una modalidad de diseño asociada al procedimiento de control experimental denominado *técnica de anidamiento*. Se trata de una estructura de investigación en la que los niveles de al menos una variable, que generalmente es de clasificación (*variable anidada*), se representan únicamente en uno de los niveles de una variable de tratamiento (*variable de anidamiento*). Por tanto, es un diseño estructuralmente incompleto en el que se generan menos condiciones experimentales que combinaciones posibles entre los factores (Arnau, 1986; Keppel, 1982; Morran, Robinson y Hulse, 1990).

Normalmente, la *variable anidada* se considera como una fuente de confundido cuyos efectos deben ser eliminados de la varianza total. En consecuencia, suele ser una *variable de naturaleza aleatoria*. Es frecuente que la variable anidada represente un grupo natural al que pertenecen los sujetos experimentales. Debido a esta circunstancia, los diseños jerárquicos son especialmente útiles en aquellos estudios en los que la pertenencia a un grupo social como, por ejemplo, el aula, el centro hospitalario en el que se está ingresado, la ciudad de residencia, etc., implica poseer una serie de características que están determinadas por el propio grupo y que influyen en la conducta de los sujetos (Arnau, 1986; Keppel, 1982). En este sentido, la utilización de los diseños jerárquicos en gran cantidad de contextos aplicados se debe, fundamentalmente, a que permiten analizar la variabilidad atribuible al efecto del factor anidado.

De hecho, los diseños jerárquicos introducen el efecto de pertenencia al grupo como fuente de variación en el modelo matemático del análisis de la varianza, controlando de esta forma la influencia que puede ejercer el grupo sobre la conducta del sujeto. En definitiva,

como señala Winer (1971), la cualidad más destacada de los diseños jerárquicos radica en que permiten eliminar una fuente de varianza sistemática secundaria, asociada a las diferencias entre los niveles del factor anidado que, en caso de no controlarse, impediría estimar con precisión los efectos del factor de anidamiento.

Con respecto a la notación que se utiliza en este tipo de diseños, cabe señalar que el factor anidado se suele representar a la izquierda de una barra diagonal o fuera de un paréntesis. A su vez, el factor de anidamiento se coloca a la derecha de la barra diagonal o dentro del paréntesis. Así, por ejemplo, en un diseño bifactorial, la notación B/A o $B(A)$ expresa que el factor B está anidado en el factor A .

8.3.2. Análisis factorial de la varianza para el diseño jerárquico

8.3.2.1. Modelo general de análisis

El modelo analítico que se utiliza, habitualmente, para llevar a cabo la **prueba de la hipótesis** en el diseño jerárquico, es el análisis factorial de la varianza. No obstante, dado que los efectos del factor anidado se hallan restringidos a un solo nivel del factor de anidamiento, la ecuación estructural del análisis de la varianza no incluye la variación atribuible a la interacción entre los factores. Si consideramos el **diseño jerárquico de dos factores**, el modelo matemático que subyace al análisis de la varianza, bajo el supuesto de la hipótesis alternativa, responde a la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \beta_{k(j)} + \varepsilon_{i(jk)} \quad (8.46)$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable de anidamiento y el k -ésimo nivel de la variable anidada.

μ = Media general de la variable dependiente.

α_j = Efecto principal asociado al j -ésimo nivel de la variable de anidamiento.

$\beta_{k(j)}$ = Efecto principal asociado al k -ésimo nivel de la variable anidada dentro del j -ésimo nivel de la variable de anidamiento.

$\varepsilon_{i(jk)}$ = Término de error o componente aleatorio del modelo.

Se asume que:

a) $\varepsilon_{i(jk)} \simeq NID(0, \sigma_\varepsilon^2)$.

b) $\beta_{k(j)} \simeq NID(0, \sigma_\beta^2)$.

Tomando como referencia el modelo matemático representado en la Ecuación (8.46), el análisis de la varianza, para el diseño jerárquico de dos factores, parte de la descomposición de la variación total en tres componentes: la variación debida al factor de anidamiento (factor A), la variación debida al factor anidado (factor B) y la variación asociada a los sujetos pertenecientes a cada uno de los niveles del factor anidado, dentro de cada uno de los niveles del factor de anidamiento, o variación residual ($\varepsilon_{i(jk)}$). A partir de las sumas de cuadrados y de sus correspondientes grados de libertad, se obtienen las medias cuadráticas del factor A y del factor B . Si se asume que el modelo matemático es un modelo mixto, en el que el factor anidado es aleatorio, el efecto de la varianza asociada al factor de anidamiento se contrasta tomando como término de error la media cuadrática del factor anidado, mientras que el término de error adecuado para contrastar el efecto del factor anidado es la media

cuadrática residual. De este modo, se estiman las razones F que permiten llevar a cabo las pruebas de la hipótesis de nulidad para cada uno de los parámetros del modelo.

Como se ha señalado previamente, la ecuación estructural del análisis de la varianza no incluye la variación atribuible a la interacción entre ambos factores, debido a que éstos no mantienen, entre sí, la relación de cruce propia de los factores que configuran un diseño factorial completo y, por tanto, dicha variación no se puede calcular. La Figura 8.2 permite ilustrar, gráficamente, la diferencia que existe entre las estructuras correspondientes a un diseño jerárquico y a un diseño factorial completo, tomando como referencia un modelo con dos factores, A y B , constando el primero de ellos de tres niveles (a_1 , a_2 y a_3) y, el segundo, de seis (b_1 , b_2 , b_3 , b_4 , b_5 y b_6).

Diseño factorial completo				
		Factor A		
		a_1	a_2	a_3
Factor B	b_1	a_1b_1	a_2b_1	a_3b_1
	b_2	a_1b_2	a_2b_2	a_3b_2
	b_3	a_1b_3	a_2b_3	a_3b_3
	b_4	a_1b_4	a_2b_4	a_3b_4
	b_5	a_1b_5	a_2b_5	a_3b_5
	b_6	a_1b_6	a_2b_6	a_3b_6

Diseño jerárquico				
		Factor A		
		a_1	a_2	a_3
Factor B	b_1	a_1b_1		
	b_2	a_1b_2		
	b_3		a_2b_3	
	b_4		a_2b_4	
	b_5			a_3b_5
	b_6			a_3b_6

Figura 8.2 Estructuras correspondientes a un diseño factorial completo y a un diseño jerárquico con dos factores, A y B , de tres y seis niveles, respectivamente.

Como cabe apreciar en la Figura 8.2, el diseño factorial completo consta de 18 condiciones experimentales, ya que cada nivel del factor A se cruza con cada uno de los niveles del factor B . Sin embargo, en el diseño jerárquico, dos de los seis niveles del factor B aparecen ligados al primer nivel del factor A , otros dos se asocian al segundo nivel de dicho factor y, los dos restantes, al tercero de sus niveles. En consecuencia, el número de condiciones experimentales es de 6, lo que implica la ausencia de determinadas combinaciones de tratamiento o la existencia de celdillas vacías en la estructura del diseño.

8.3.2.2. Consideraciones importantes con respecto a la selección del término de error para contrastar el efecto de cada uno de los factores incluidos en el diseño

Como en todo modelo de análisis de la varianza, en el caso del diseño jerárquico, el contraste de la hipótesis a través de la razón F también implica la comparación entre la varianza intergrupos y la varianza intragrupo. Sin embargo, como afirman Keppel (1982) y Maxwell y Delaney (1990), el término de error adecuado, para llevar a cabo dicho contraste, depende del tipo de modelo estadístico que se asuma. Ya hemos visto en el capítulo anterior, que el modelo de efectos fijos se asocia, habitualmente, a los diseños factoriales completamente aleatorios. En este caso, la media cuadrática del error constituye el término de error (denominador de la razón F) adecuado para estimar los efectos correspondientes a cada una de las fuentes de variación del diseño. No obstante, en los diseños jerárquicos, las *variables anidadas* se suelen considerar *factores de efectos aleatorios* y, en consecuencia, todos los

contrastes de las hipótesis de nulidad asociadas a los parámetros de la ecuación estructural no asumen el mismo término de error.

El caso más común de diseño jerárquico de dos factores asume que el modelo estadístico es un modelo mixto, siendo la variable de anidamiento de efectos fijos y la variable anidada de efectos aleatorios. En este tipo de modelo, que corresponde a la ecuación estructural planteada en la Fórmula (8.46), el término de error adecuado para el factor de anidamiento es la media cuadrática del factor anidado. Por otra parte, el efecto del factor anidado se contrasta tomando, como término de error, la media cuadrática del error. Como principio general, válido para cualquier diseño jerárquico, Pascual, García y Frías (1995) plantean que la variabilidad asociada a un factor anidado de efectos aleatorios se transmite a todas las fuentes de variación que se encuentran por encima de dicho factor. En consecuencia, el término de error adecuado para tales fuentes de variación es la media cuadrática del factor anidado.

Como cabe deducir de este principio, cuando el factor anidado asume un modelo de efectos aleatorios, el contraste de la hipótesis implica un menor número de grados de libertad para el término de error utilizado en la razón F , lo que disminuye la sensibilidad de la prueba estadística (Keppel, 1982). No obstante, dicho problema se puede solventar aumentando el número de niveles del factor anidado. Por otra parte, en caso de que el factor anidado no ejerza una influencia significativa sobre la variable de respuesta, es posible agrupar su variabilidad y la del error en un solo componente. De esta manera también se incrementan los grados de libertad del denominador de la razón F , obteniéndose un término de error más estable para contrastar los efectos de los tratamientos (Arnau, 1994; Keppel, 1982).

En la Figura 8.3 se proporcionan las pautas que debe seguir el investigador para seleccionar adecuadamente el término de error en el análisis de la varianza, en función del modelo estadístico (de efectos fijos o de efectos aleatorios) asumido para cada uno de los factores que configuran el diseño.

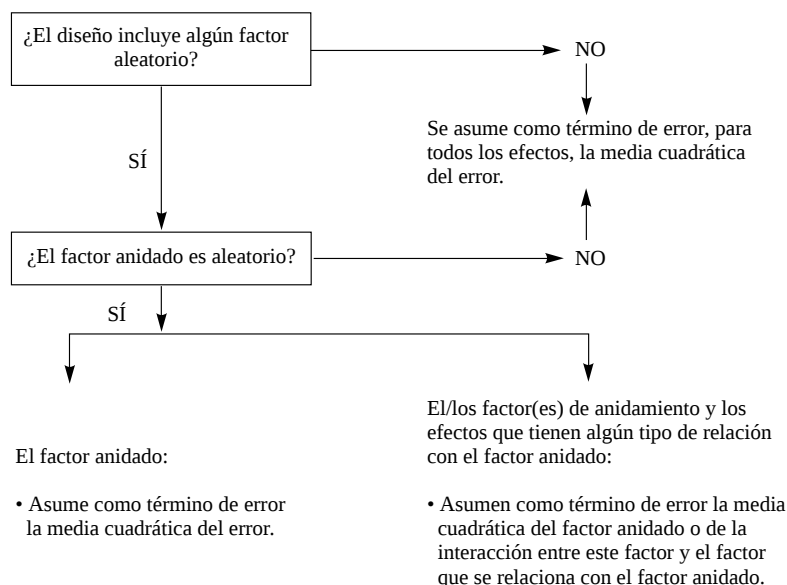


Figura 8.3 Pautas para la adecuada selección del término de error en la prueba de la hipótesis de nulidad para cada uno de los parámetros del modelo correspondiente al diseño jerárquico (adaptada de Maxwell y Delaney, 1990).

A fin de ilustrar el proceso de selección del término de error en diferentes modelos de diseño jerárquico, proporcionamos las representaciones correspondientes a una serie de diseños jerárquicos de dos y de tres factores, así como diversas tablas que resultan muy útiles para adoptar la decisión correcta con respecto a la selección del término de error, asociado a cada uno de los efectos incluidos en el diseño, en función de la naturaleza de sus factores. A efectos meramente didácticos, en todos los casos se presupone que cada factor consta únicamente de dos niveles.

• **Diseño jerárquico de dos factores con un factor de anidamiento y un factor anidado**

En la Figura 8.4 se representa la estructura correspondiente a este tipo de diseño. A su vez, en la Tabla 8.16 se proporcionan los términos de error adecuados para contrastar los efectos incluidos en el mismo, en función del modelo estadístico asumido para cada uno de sus factores.

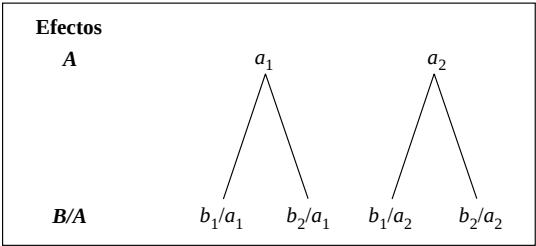


Figura 8.4 Estructura correspondiente a un diseño jerárquico de dos factores con un factor de anidamiento (factor A) y un factor anidado (factor B/A).

TABLA 8.16 Representación de los términos de error adecuados para contrastar los efectos incluidos en un diseño jerárquico de dos factores con un factor de anidamiento (factor A) y un factor anidado (factor B/A), en función del modelo estadístico asumido para cada uno de tales factores

Efectos	Modelo de efectos fijos	Modelo de efectos mixtos		Modelo de efectos aleatorios
	A B/A	A fijo B/A aleatorio	A aleatorio B/A fijo	A B/A
A	MC_{error}	$MC_{B/A}$	MC_{error}	$MC_{B/A}$
B/A	MC_{error}	MC_{error}	MC_{error}	MC_{error}

Cuando el factor anidado o ambos factores son de efectos fijos, el término de error adecuado, para contrastar el efecto de cualquiera de los dos factores que configuran el diseño, es la media cuadrática del error. No obstante, cuando la variable anidada o ambas variables son de naturaleza aleatoria, el efecto de la variable de anidamiento se contrasta tomando como término de error la media cuadrática del factor anidado, mientras que el componente de error apropiado para contrastar el efecto del factor anidado es la media cuadrática del error (véase la Figura 8.3).

• **Diseño jerárquico de tres factores con un factor de anidamiento, un factor anidado y un factor que mantiene una relación de cruce con el factor anidado**

En la Figura 8.5 se representa la estructura correspondiente a este tipo de diseño. A su vez, en la Tabla 8.17 se muestran los términos de error adecuados para contrastar los efectos incluidos en el mismo, en función del modelo estadístico asumido para cada uno de sus factores.

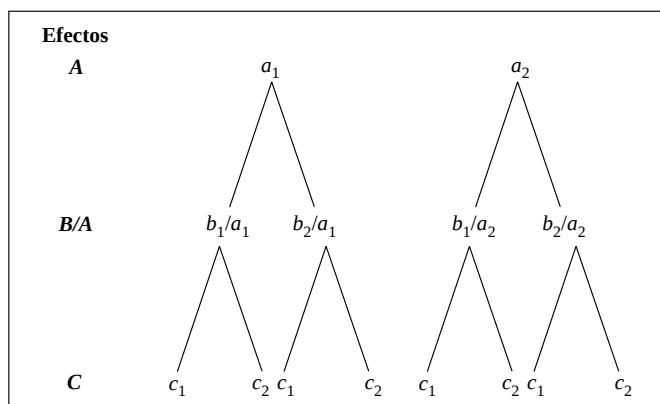


Figura 8.5 Estructura correspondiente a un diseño jerárquico de tres factores con un factor de anidamiento (factor A), un factor anidado (factor B/A) y un factor que mantiene una relación de cruce con el factor anidado (factor C).

TABLA 8.17 Representación de los términos de error adecuados para contrastar los efectos incluidos en un diseño jerárquico de tres factores con un factor de anidamiento (factor A), un factor anidado (factor B/A) y un factor que mantiene una relación de cruce con el factor anidado (factor C), en función del modelo estadístico asumido para cada uno de tales factores

Efectos	Modelo de efectos fijos	Modelo de efectos mixtos		Modelo de efectos aleatorios
	A B/A C	B/A aleatorio	B/A fijo	A B/A C
A	MC_{error}	$MC_{B/A}$	MC_{error}	$MC_{B/A}$
B/A	MC_{error}	MC_{error}	MC_{error}	MC_{error}
C	MC_{error}	$MC_{B/A \times C}$	MC_{error}	$MC_{B/A \times C}$
$A \times C$	MC_{error}	$MC_{B/A \times C}$	MC_{error}	$MC_{B/A \times C}$
$B/A \times C$	MC_{error}	MC_{error}	MC_{error}	MC_{error}

En primer lugar, cabe señalar que la naturaleza de los factores no anidados (factor *A* y factor *C*) no condiciona los términos de error que deben utilizarse para contrastar los diferentes parámetros del modelo. Por otra parte, cuando el factor anidado o los tres factores que configuran el diseño son de efectos fijos, el término de error adecuado para contrastar cualquiera de los efectos incluidos en el mismo es la media cuadrática del error. Sin embargo, cuando la variable anidada o las tres variables son de naturaleza aleatoria, la variable de anidamiento asume como término de error la media cuadrática de la variable anidada. A su vez, aunque el factor *C* no es un factor de anidamiento, mantiene una relación de cruce con un factor anidado de naturaleza aleatoria y, en consecuencia, tanto su efecto como el de la interacción $A \times C$ se hallan bajo el influjo de los diferentes niveles de la variable anidada. Debido a esta circunstancia, el término de error adecuado para contrastar, tanto el efecto de *C* como el de $A \times C$ es la media cuadrática de la interacción entre el factor anidado y el factor *C* ($MC_{B/A \times C}$). Por último, el componente de error apropiado para contrastar el efecto del factor anidado, u otro efecto cualquiera que implique a dicho factor, en nuestro caso el debido a la interacción $B/A \times C$, es la media cuadrática del error.

• **Diseño jerárquico de tres factores con un factor de anidamiento y dos factores anidados**

En la Figura 8.6 se representa la estructura correspondiente a este tipo de diseño. A su vez, en la Tabla 8.18 se proporcionan los términos de error adecuados para contrastar los efectos incluidos en el mismo, en función del modelo estadístico asumido para cada uno de sus factores.

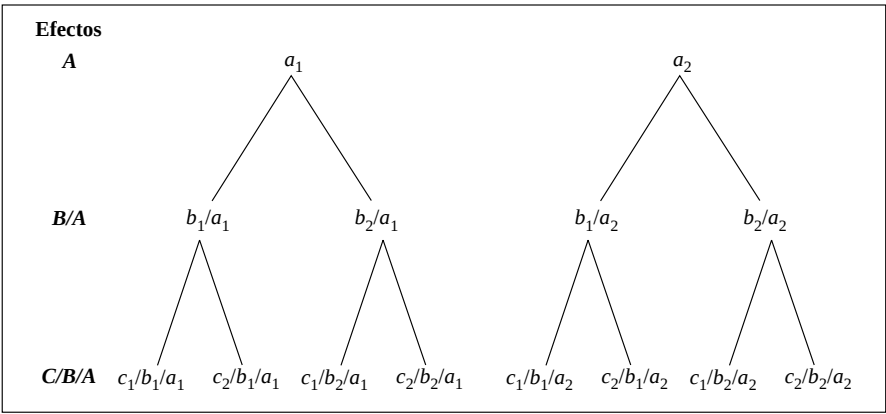


Figura 8.6 Estructura correspondiente a un diseño jerárquico de tres factores con un factor de anidamiento (factor *A*), y dos factores anidados (factor *B/A* y factor *C/B/A*).

Cuando ambos factores anidados o los tres factores que configuran el diseño son de efectos fijos, el término de error adecuado para contrastar cualquiera de los efectos incluidos en el mismo, es la media cuadrática del error. No obstante, cuando las dos variables anidadas o las tres variables son de naturaleza aleatoria, la variabilidad asociada a cada una de tales variables se transmite a todas las fuentes de variación que se encuentran encima de ellas. En consecuencia, el efecto de cada factor se contrasta tomando como término de error la media

TABLA 8.18 Representación de los términos de error adecuados para contrastar los efectos incluidos en un diseño jerárquico de tres factores con un factor de anidamiento (factor A), y dos factores anidados (factor B/A y factor C/B/A), en función del modelo estadístico asumido para cada uno de tales factores

Efectos	Modelo de efectos fijos	Modelo de efectos mixtos		Modelo de efectos aleatorios
	A B/A C/B/A	A fijo B/A aleatorio C/B/A aleatorio	A aleatorio B/A fijo C/B/A fijo	A B/A C/B/A
A	MC_{error}	$MC_{B/A}$	MC_{error}	$MC_{B/A}$
B/A	MC_{error}	$MC_{C/B/A}$	MC_{error}	$MC_{C/B/A}$
C/B/A	MC_{error}	MC_{error}	MC_{error}	MC_{error}

cuadrática del efecto que se encuentra inmediatamente debajo de él, en la estructura del diseño ($MC_{B/A}$ para el factor A y $MC_{C/B/A}$ para el factor B/A). Por último, cabe señalar que el componente de error apropiado para contrastar el efecto del factor ubicado en la parte más baja, dentro de la jerarquía del diseño, a saber, del factor C/B/A es la media cuadrática del error.

Dado que el objetivo del presente subapartado consistía, básicamente, en subrayar el hecho de que la selección del término de error adecuado para contrastar los efectos de los diferentes factores incluidos en un diseño jerárquico, requiere tomar en consideración el modelo estadístico asociado a cada uno de tales factores, no nos extenderemos más en esta cuestión. No obstante, remitimos al lector interesado en dicha problemática a los textos de Keppel (1982) y de Pascual, García y Frías (1995).

8.3.2.3. Ejemplo práctico

Supongamos que un asistente social desea examinar la influencia que ejercen *tres tipos de programas para el aprendizaje de conductas de aseo personal* (factor A) en la *cantidad de conductas de este tipo que desarrollan los ancianos internados en residencias o centros para la tercera edad*. A fin de dilucidar tal cuestión, lleva a cabo un experimento con una muestra de 18 ancianos pertenecientes a *seis residencias diferentes* (factor B/A). Tras implantar los programas, los responsables de cada centro registran la cantidad de conductas de aseo personal que desarrolla diariamente cada uno de los sujetos experimentales, durante un determinado período temporal. En la Tabla 8.19 pueden observarse los resultados obtenidos en la investigación.

En el presente diseño, calcularemos las sumas de cuadrados mediante los dos procedimientos que hemos empleado con los diseños abordados en los capítulos anteriores. A partir de tales sumas de cuadrados, estimaremos las varianzas y las razones F asociadas a los distintos parámetros de la ecuación estructural del ANOVA.

TABLA 8.19 Matriz de datos del experimento

A (Programa)	a ₁ (Programa 1)		a ₂ (Programa 2)		a ₃ (Programa 3)		
B/A (Centro/ Programa)	b ₁ /a ₁ (Centro 1/ Prog. 1)	b ₂ /a ₁ (Centro 2/ Prog. 1)	b ₃ /a ₂ (Centro 3/ Prog. 2)	b ₄ /a ₂ (Centro 4/ Prog. 2)	b ₅ /a ₃ (Centro 5/ Prog. 3)	b ₆ /a ₃ (Centro 6/ Prog. 3)	
	12	17	11	5	4	12	
	12	13	15	9	1	8	
	13	20	12	12	6	10	
Sumatorios y medias marginales	$\Sigma Y_{11}=37$ $\bar{Y}_{11}=12,33$	$\Sigma Y_{12}=50$ $\bar{Y}_{12}=16,67$	$\Sigma Y_{23}=38$ $\bar{Y}_{23}=12,67$	$\Sigma Y_{24}=26$ $\bar{Y}_{24}=8,67$	$\Sigma Y_{35}=11$ $\bar{Y}_{35}=3,67$	$\Sigma Y_{36}=30$ $\bar{Y}_{36}=10$	$\Sigma Y_{..}=192$ $\bar{Y}_{..}=10,67$

Procedimiento 1

En la Tabla 8.20 presentamos las fórmulas necesarias para el cálculo de las sumas cuadráticas.

TABLA 8.20 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño jerárquico de dos factores, en el que el factor anidado es aleatorio

Variabilidad intergrupos	$SC_{\text{intergrupos}} = \left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C$ (8.48)
Efecto del factor de anidamiento o factor A	$SCA = \left[\frac{1}{nb} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C$ (8.49) donde b = cantidad de centros en los que se implanta cada programa
Efecto del factor anidado o factor B/A	$SCB/A = SC_{\text{intergrupos}} - SCA$ (8.50)
Variabilidad total	$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$ (8.51)
Variabilidad residual o del error	$SCE \text{ o } SCS/B/A = SCT - SC_{\text{intergrupos}}$ (8.52)
C	$C = \frac{1}{abn} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2$ (8.53)

Tras obtener el valor de C, llevaremos a cabo el cálculo de las sumas de cuadrados.

$$C = \frac{1}{3 \cdot 2 \cdot 3} (12 + 12 + 13 + 17 + 13 + 20 + 11 + 15 + 12 + 5 + 9 + 12 + 4 + 1 + 6 + 12 + 8 + 10)^2 = \frac{(192)^2}{18} = 2.048$$

Una vez obtenido este valor, procedemos a calcular la $SC_{\text{intergrupos}}$, la SCA , la SCB/A , la SCT y la $SCS/B/A$, respectivamente.

$$SC_{\text{intergrupos}} = \frac{1}{3} [(37)^2 + (50)^2 + (38)^2 + (26)^2 + (11)^2 + (30)^2] - 2.048 = \\ = 2.336,67 - 2.048 = 288,67$$

$$SCA = \frac{1}{3 \cdot 2} [(37 + 50)^2 + (38 + 26)^2 + (11 + 30)^2] - 2.048 = 2.224,33 - 2.048 = 176,33$$

$$SCB/A = 288,67 - 176,33 = 112,34$$

$$SCT = \left[\begin{array}{l} (12)^2 + (12)^2 + (13)^2 + (17)^2 + (13)^2 + (20)^2 + (11)^2 + (15)^2 + (12)^2 + \\ + (5)^2 + (9)^2 + (12)^2 + (4)^2 + (1)^2 + (6)^2 + (12)^2 + (8)^2 + (10)^2 \end{array} \right] - 2.048 = \\ = 2.416 - 2.048 = 368$$

$$SCS/B/A = 368 - 288,67 = 79,33$$

Procedimiento 2: Desarrollo mediante vectores

En primer lugar, hallaremos los vectores asociados a cada uno de los parámetros de la ecuación estructural del ANOVA correspondiente al diseño jerárquico de dos factores (fórmula 8.46) y, posteriormente, calcularemos las sumas de cuadrados de los diferentes efectos.

VECTOR Y

Como ya es sabido, este vector es el *vector column* de todas las *puntuaciones directas*.

$$Y = \{y\} = \begin{bmatrix} 12 \\ 12 \\ 13 \\ 17 \\ 13 \\ 20 \\ 11 \\ 15 \\ 12 \\ 5 \\ 9 \\ 12 \\ 4 \\ 1 \\ 6 \\ 12 \\ 8 \\ 10 \end{bmatrix} \quad (8.54)$$

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada una de las categorías del factor de anidamiento o factor A para estimar, a partir de tales medias, los parámetros asociados al efecto de dicho factor, α_j .

$$\mu = \frac{1}{nab} \sum_i \sum_j \sum_k Y_{ijk} \quad (8.55)$$

$$\mu = \frac{1}{3 \cdot 3 \cdot 2} [12 + 12 + 13 + 17 + 13 + 20 + 11 + 15 + 12 + 5 + 9 + 12 + 4 + 1 + 6 + 12 + 8 + 10] = \frac{192}{18} = 10,67$$

Promedios de los niveles del factor de anidamiento o factor A:

$$\alpha_j = \mu_{j.} - \mu \quad (8.56)$$

$$\mu_{j.} = \frac{1}{nb} \sum_i \sum_k Y_{ijk}$$

respectivamente:

$$\mu_{1.} = \frac{1}{3 \cdot 2} [12 + 12 + 13 + 17 + 13 + 20] = 14,5$$

$$\mu_{2.} = \frac{1}{3 \cdot 2} [11 + 15 + 12 + 5 + 9 + 12] = 10,67$$

$$\mu_{3.} = \frac{1}{3 \cdot 2} [4 + 1 + 6 + 12 + 8 + 10] = 6,83$$

Por tanto:

$$\alpha_1 = \mu_{1.} - \mu = 14,5 - 10,67 = 3,83$$

$$\alpha_2 = \mu_{2.} - \mu = 10,67 - 10,67 = 0$$

$$\alpha_3 = \mu_{3.} - \mu = 6,83 - 10,67 = -3,84$$

El *vector* A adopta los siguientes valores:

$$A = \{\alpha\} = \begin{bmatrix} 3,83 \\ 3,83 \\ 3,83 \\ 3,83 \\ 3,83 \\ 3,83 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -3,84 \\ -3,84 \\ -3,84 \\ -3,84 \\ -3,84 \\ -3,84 \end{bmatrix} \quad (8.57)$$

VECTOR B/A

A fin de obtener la variabilidad correspondiente al factor anidado o factor B/A , estimaremos los parámetros asociados a dicho factor, $\beta_{k/j}$. Para ello, calcularemos la media de cada combinación entre cada uno de los niveles del factor A y del factor B , $\mu_{k/j}$, y restaremos a cada uno de tales valores el promedio de la categoría del factor A en el que se lleva a cabo el anidamiento, μ_j .

$$\beta_{k/j} = \mu_{k/j} - \mu_j. \quad (8.58)$$

$$\mu_{k/j} = \frac{1}{n} \sum_i Y_{ijk}$$

Promedios que corresponden a las combinaciones entre los niveles del factor de anidamiento y del factor anidado:

$$\mu_{1/1} = \frac{1}{3} [12 + 12 + 13] = 12,33$$

$$\mu_{2/1} = \frac{1}{3} [17 + 13 + 20] = 16,67$$

$$\mu_{3/2} = \frac{1}{3} [11 + 15 + 12] = 12,67$$

$$\mu_{4/2} = \frac{1}{3} [5 + 9 + 12] = 8,67$$

$$\mu_{5/3} = \frac{1}{3} [4 + 1 + 6] = 3,67$$

$$\mu_{6/3} = \frac{1}{3} [12 + 8 + 10] = 10$$

Por tanto:

$$\beta_{1/1} = \mu_{1/1} - \mu_{1.} = 12,33 - 14,5 = -2,17$$

$$\beta_{2/1} = \mu_{2/1} - \mu_{1.} = 16,67 - 14,5 = 2,17$$

$$\beta_{3/2} = \mu_{3/2} - \mu_{2.} = 12,67 - 10,67 = 2$$

$$\beta_{4/2} = \mu_{4/2} - \mu_{2.} = 8,67 - 10,67 = -2$$

$$\beta_{5/3} = \mu_{5/3} - \mu_{3.} = 3,67 - 6,83 = -3,16$$

$$\beta_{6/3} = \mu_{6/3} - \mu_{3.} = 10 - 6,83 = 3,17$$

El vector B/A adopta los siguientes valores:

$$B/A = \{\beta/\alpha\} = \begin{bmatrix} -2,17 \\ -2,17 \\ -2,17 \\ 2,17 \\ 2,17 \\ 2,17 \\ 2 \\ 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ -3,16 \\ -3,16 \\ -3,16 \\ 3,17 \\ 3,17 \\ 3,17 \end{bmatrix} \quad (8.59)$$

VECTOR E , CÁLCULO DE LOS RESIDUALES O ERRORES

Partiendo de la Fórmula (8.46):

$$\varepsilon_{i(jk)} = Y_{ijk} - \mu - \alpha_j - \beta_{k/j}$$

$$\varepsilon_{i(11)} = Y_{i11} - \mu - \alpha_1 - \beta_{1/1} = Y_{i11} - 10,67 - 3,83 - (-2,17) = Y_{i11} - 12,33$$

Sustituyendo las puntuaciones Y_{i11} por sus correspondientes valores:

$$\varepsilon_{1(11)} = 12 - 12,33 = -0,33$$

$$\varepsilon_{2(11)} = 12 - 12,33 = -0,33$$

$$\varepsilon_{3(11)} = 13 - 12,33 = 0,67$$

$$\varepsilon_{i(12)} = Y_{i12} - \mu - \alpha_1 - \beta_{2/1} = Y_{i12} - 10,67 - 3,83 - 2,17 = Y_{i12} - 16,67$$

Sustituyendo las puntuaciones Y_{i12} por sus correspondientes valores:

$$\varepsilon_{1(12)} = 17 - 16,67 = 0,33$$

$$\varepsilon_{2(12)} = 13 - 16,67 = -3,67$$

$$\varepsilon_{3(12)} = 20 - 16,67 = 3,33$$

$$\varepsilon_{i(23)} = Y_{i23} - \mu - \alpha_2 - \beta_{3/2} = Y_{i23} - 10,67 - 0 - 2 = Y_{i23} - 12,67$$

Sustituyendo las puntuaciones Y_{i23} por sus correspondientes valores:

$$\varepsilon_{1(23)} = 11 - 12,67 = -1,67$$

$$\varepsilon_{2(23)} = 15 - 12,67 = 2,33$$

$$\varepsilon_{3(23)} = 12 - 12,67 = -0,67$$

$$\varepsilon_{i(24)} = Y_{i24} - \mu - \alpha_2 - \beta_{4/2} = Y_{i24} - 10,67 - 0 - (-2) = Y_{i24} - 8,67$$

Sustituyendo las puntuaciones Y_{i24} por sus correspondientes valores:

$$\varepsilon_{1(24)} = 5 - 8,67 = -3,67$$

$$\varepsilon_{2(24)} = 9 - 8,67 = 0,33$$

$$\varepsilon_{3(24)} = 12 - 8,67 = 3,33$$

$$\varepsilon_{i(35)} = Y_{i35} - \mu - \alpha_3 - \beta_{5/3} = Y_{i35} - 10,67 - (-3,84) - (-3,16) = Y_{i35} - 3,67$$

Sustituyendo las puntuaciones Y_{i35} por sus correspondientes valores:

$$\varepsilon_{1(35)} = 4 - 3,67 = 0,33$$

$$\varepsilon_{2(35)} = 1 - 3,67 = -2,67$$

$$\varepsilon_{3(35)} = 6 - 3,67 = 2,33$$

$$\varepsilon_{i(36)} = Y_{i36} - \mu - \alpha_3 - \beta_{6/3} = Y_{i36} - 10,67 - (-3,84) - 3,17 = Y_{i36} - 10$$

Sustituyendo las puntuaciones Y_{i36} por sus correspondientes valores:

$$\varepsilon_{1(36)} = 12 - 10 = 2$$

$$\varepsilon_{2(36)} = 8 - 10 = -2$$

$$\varepsilon_{3(36)} = 10 - 10 = 0$$

Por tanto, el *vector E* adopta los siguientes valores:

$$E = \{\varepsilon\} = \begin{bmatrix} -0,33 \\ -0,33 \\ 0,67 \\ 0,33 \\ -3,67 \\ 3,33 \\ -1,67 \\ 2,33 \\ -0,67 \\ -3,67 \\ 0,33 \\ 0,33 \\ 0,33 \\ -2,67 \\ 2,33 \\ 2 \\ -2 \\ 0 \end{bmatrix} \quad (8.60)$$

Llegados a este punto, podemos calcular la *SCA*, la *SCB/A* y la *SCS/B/A* aplicando las Fórmulas (8.61), (8.62) y (8.63), respectivamente, o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos.

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR DE ANIDAMIENTO O FACTOR A (SCA):

$$SCA = nb \sum_j \alpha_j^2 = 3 \cdot 2[(3,83)^2 + (0)^2 + (-3,84)^2] = 176,48 \quad (8.61)$$

SUMA CUADRÁTICA CORRESPONDIENTE AL FACTOR ANIDADO O FACTOR B (SCB/A):

$$SCB/A = n \sum_j \sum_k \beta_{k/j}^2 = 3[(-2,17)^2 + (2,17)^2 + (2)^2 + (-2)^2 + (-3,16)^2 + (3,16)^2] = 112,17 \quad (8.62)$$

SUMA CUADRÁTICA CORRESPONDIENTE A LOS RESIDUALES O ERRORES (SCS/B/A):

$$SCS/B/A = \sum_i \sum_j \sum_k \varepsilon_{ijk}^2 = (-0,33)^2 + (-0,33)^2 + (0,67)^2 + (0,33)^2 + (-3,67)^2 + (3,33)^2 + (-1,67)^2 + (2,33)^2 + (-0,67)^2 + (-3,67)^2 + (0,33)^2 + (3,33)^2 + (0,33)^2 + (-2,67)^2 + (2,33)^2 + (2)^2 + (-2)^2 + (0)^2 = 79,33 \quad (8.63)$$

Multiplicando cada vector por su traspuesto obtenemos los mismos resultados.

$$SCA = \{\alpha\}^T \{\alpha\} = 176,48 \quad (8.64)$$

$$SCB/A = \{\beta/\alpha\}^T \{\beta/\alpha\} = 112,17 \quad (8.65)$$

$$SCS/B/A = \{\varepsilon\}^T \{\varepsilon\} = 79,33 \quad (8.66)$$

No debemos olvidar que la $SCT = SCA + SCB/A + SCS/B/A = 367,98$.

Tras obtener las sumas de cuadrados mediante diferentes procedimientos, podemos elaborar la tabla del análisis de la varianza.

TABLA 8.21 Análisis factorial de la varianza para un diseño jerárquico de dos factores, en el que el factor anidado es aleatorio: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A	$SCA = 176,48$	$a - 1 = 2$	$MCA = 88,24$	$F_A = \frac{MCA}{MCB/A} = 2,36$
Factor B/A	$SCB/A = 112,17$	$a(b/a - 1) = 3$	$MCB/A = 37,39$	$F_{B/A} = \frac{MCB/A}{MCS/B/A} = 5,66$
Error (S/B/A)	$SCS/B/A = 79,33$	$ab(n - 1) = 12$	$MCS/B/A = 6,61$	
Total	$SCT = 367,98$	$abn - 1 = 17$		

En función de los resultados obtenidos en el ANOVA cabe concluir que el factor de anidamiento, es decir, el *tipo de programa que se les imparte a los ancianos* no ejerce una influencia estadísticamente significativa sobre la *cantidad de conductas de aseo que éstos desarrollan diariamente* ($F_{0,95; 2,3} = 9,55 > F_{obs.} = 2,36$). Sin embargo, la variable anidada o el *centro en el que se encuentran internados los ancianos* explica una proporción estadísticamente significativa de la varianza total ($F_{0,95; 3,12} = 3,49 < F_{obs.} = 5,66$). En consecuencia, cabe afirmar que, en el presente estudio, la utilización de un diseño jerárquico resulta muy adecuada, ya que permite eliminar una fuente de varianza sistemática secundaria asociada a

las diferencias entre los niveles del factor anidado, y de este modo incrementar la potencia del diseño.

8.3.2.4. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la(s) variable(s) dependiente(s) y los factores, especificando en su lugar correspondiente si tales factores son fijos o aleatorios (en nuestro ejemplo, el factor de anidamiento o factor A es fijo y el factor anidado o factor B es aleatorio).



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar «estadísticos descriptivos»*, *«estimaciones del tamaño del efecto»*, *«potencia observada»* y *«pruebas de homogeneidad»*).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería por defecto, utilizando la opción **Pegar**:

UNIANOVA

```
conducta BY programa centro
/RANDOM = centro
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(programa)
/EMMEANS = TABLES(centro)
/PRINT = DESCRIPTIVE ETASQ OPOWER HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = programa centro programa*centro.
```

A continuación, debemos especificar el tipo de diseño que, por defecto, el programa concibe como un diseño factorial, indicando que se trata de un diseño jerárquico de dos factores. Esta especificación únicamente puede realizarse modificando la sintaxis. Para ello, en el subcomando /DESIGN debemos escribir lo siguiente: «programa centro(programa)» o bien «programa centro within programa», lo cual indica que el diseño incluye el efecto principal del factor «programa» y el efecto del factor «centro» anidado en el factor «programa»¹. Una vez realizada esta corrección, la sintaxis correspondiente a nuestro ejemplo de diseño jerárquico de dos factores, en el que el factor anidado es aleatorio, sería:

UNIANOVA

```
conducta BY programa centro
/RANDOM = centro
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(programa)
/EMMEANS = TABLES(centro)
/PRINT = DESCRIPTIVE ETASQ OPOWER HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = programa centro(programa).
```

- Una vez ejecutada la sintaxis, los resultados obtenidos son:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Programa	1,00	Programa 1	6
	2,00	Programa 2	6
	3,00	Programa 3	6
Centro	1,00	Centro 1	3
	2,00	Centro 2	3
	3,00	Centro 3	3
	4,00	Centro 4	3
	5,00	Centro 5	3
	6,00	Centro 6	3

¹ Para indicar que el factor B está anidado en A, la sintaxis es B(A) o bien B within A. Si existen más factores anidados, por ejemplo, B anidado en C y A anidado en B(C), deberíamos indicar: A(B(C)). Para más información al respecto, puede consultarse la *Guía de Sintaxis* del programa SPSS 10.0.

Estadísticos descriptivos**Variable dependiente: Conducta de aseo**

Programa	Centro	Media	Desv. típ.	N
Programa 1	Centro 1	12,3333	0,5774	3
	Centro 2	16,6667	3,5119	3
	Total	14,5000	3,2711	6
Programa 2	Centro 3	12,6667	2,0817	3
	Centro 4	8,6667	3,5119	3
	Total	10,6667	3,3862	6
Programa 3	Centro 5	3,6667	2,5166	3
	Centro 6	10,0000	2,0000	3
	Total	6,8333	4,0208	6
Total	Centro 1	12,3333	0,5774	3
	Centro 2	16,6667	3,5119	3
	Centro 3	12,6667	2,0817	3
	Centro 4	8,6667	3,5119	3
	Centro 5	3,6667	2,5166	3
	Centro 6	10,0000	2,0000	3
	Total	10,6667	4,6526	18

Contraste de Levene sobre la igualdad de las varianzas error^a**Variable dependiente: Conducta de aseo**

F	gl1	gl2	Sig.
0,982	5	12	0,467

Contrasta la hipótesis nula de que la varianza error de la variable dependiente es igual a lo largo de todos los grupos.

^a Diseño: Intercept + PROGRAMA + CENTRO(PROGRAMA)

Pruebas de los efectos intersujetos**Variable dependiente: Conducta de aseo**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Intercept								
Hipótesis	2.048,000	1	2.048,000	54,69	0,005	0,948	54,694	0,995
Error	112,333	3	37,444 ^b					
PROGRAMA								
Hipótesis	176,333	2	88,167	2,355	0,243	0,611	4,709	0,209
Error	112,333	3	37,444 ^b					
CENTRO(PROGRAMA)								
Hipótesis	112,333	3	37,444	5,664	0,012	0,586	16,992	0,846
Error	79,333	12	6,611 ^c					

^a Calculado con alfa = 0,05.

^b MS(CENTRO(PROGRAMA)).

^c MS(Error).

Media cuadrática esperada ^{a, b}

Fuente	Componente de la varianza		
	Var(CENTRO(PROGRAMA))	Var(Error)	Término cuadrático
Intercept	3.000	1.000	Intercept, PROGRAMA PROGRAMA
PROGRAMA	3,000	1,000	
CENTRO(PROGRAMA)	3,000	1,000	
Error	0,000	1,000	

^a Para cada fuente, la media cuadrática esperada es igual a la suma de los coeficientes de las casillas por las componentes de la varianza, más un término cuadrático que incluye los efectos de la casilla Término cuadrático.

^b Las medias cuadráticas esperadas se basan en la suma de cuadrados tipo III.

Medias marginales estimadas

1. Programa

Variable dependiente: Conducta de aseo

Programa	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Programa 1	14,500 ^a	1,050	12,213	16,787
Programa 2	10,667 ^a	1,050	8,380	12,954
Programa 3	6,833 ^a	1,050	4,546	9,120

^a Basada en la media marginal poblacional modificada.

2. Centro

Variable dependiente: Conducta de aseo

Centro	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Centro 1	12,333 ^a	1,484	9,099	15,568
Centro 2	16,667 ^a	1,484	13,432	19,901
Centro 3	12,667 ^a	1,484	9,432	15,901
Centro 4	8,667 ^a	1,484	5,432	11,901
Centro 5	3,667 ^a	1,484	0,432	6,901
Centro 6	10,000 ^a	1,484	6,766	13,234

^a Basada en la media marginal poblacional modificada.

8.4. DISEÑOS CON COVARIABLES

8.4.1. Características generales del diseño con covariables

Cuando el influjo de una o más variables extrañas no se puede eliminar mediante procedimientos de control experimental, cabe la posibilidad de controlarlo realizando determinados

ajustes en el análisis estadístico. Si las variables extrañas son cuantitativas, el mejor procedimiento de ajuste es el *análisis de la covarianza (ANCOVA)*. Los diseños que incorporan en su estructura esta técnica de control reciben el nombre de *diseños de covarianza* o *diseños con covariables*. Se trata de diseños que incluyen variables cuantitativas y cualitativas en su componente sistemático, y cuyo objetivo fundamental consiste en eliminar la influencia que ejercen, sobre la variable dependiente, las variables cuantitativas o covariables. Así, la característica principal de este tipo de modelos es la aplicación de un control estadístico, que permite ajustar las puntuaciones de la variable dependiente y evitar el sesgo sistemático que pueden ocasionar, en los resultados, una o más covariables. En consecuencia, son diseños que proporcionan una estimación más precisa de la incidencia de los tratamientos administrados (Elashoff, 1969; Huitema, 1980; Stevens, 1992).

Como afirman Pascual, Frías y García (1996), el análisis de la covarianza es una técnica general de análisis estadístico que puede aplicarse tanto a los datos procedentes de diseños experimentales como cuasiexperimentales, e independientemente del modelo de diseño (univariado o multivariado, intersujetos o intrasujeto, unifactorial o factorial, etc.) con el que se trabaje. No obstante, y aunque en todos los casos se persiga el objetivo de controlar la variabilidad no deseada por el investigador, el análisis de la covarianza cumple funciones distintas dependiendo de la naturaleza de la investigación en la que se aplique. Así, mientras en la investigación experimental dicha técnica permite reducir el error aleatorio e incrementar la precisión en la estimación de los efectos experimentales, en la investigación cuasiexperimental posibilita eliminar el sesgo sistemático derivado de las diferencias existentes a priori entre los grupos de tratamiento, las cuales son consecuencia de la ausencia de aleatorización en la asignación de los sujetos a los tratamientos, propia de los diseños cuasiexperimentales. Hemos de señalar que, dado que en el presente texto se abordan los diseños experimentales, desarrollaremos el análisis de la covarianza en el marco de la investigación experimental.

En palabras de Fisher (1925), el *análisis de la covarianza* combina las ventajas e integra en un solo procedimiento dos métodos de gran aplicabilidad: el *análisis de la regresión* y el *análisis de la varianza*. En el análisis de la covarianza se incluyen tres tipos de variables: la(s) *variable(s) independiente(s)* o variable(s) cuyos efectos se quieren estimar, la(s) *variable(s) dependiente(s)* o respuesta(s) que se espera(n) encontrar como consecuencia de la aplicación de los tratamientos experimentales y la(s) *covariable(s)* o variable(s) extraña(s) no sometida(s) a investigación y cuyos efectos se pretenden controlar mediante procedimientos de ajuste estadístico (Pascual, 1995b). Como señala Huitema (1980), el *análisis de la covarianza* se realiza a partir del *análisis de la regresión*. En primer lugar, se efectúa un análisis de regresión de la *variable dependiente* sobre la(s) *variable(s) de tratamiento* y sobre la(s) *covariable(s)*. De esta manera, se obtiene el *coeficiente de determinación múltiple* que representa la proporción de varianza de la variable dependiente explicada por la(s) variable(s) de tratamiento y por la(s) covariable(s). A continuación, se regresiona la *variable dependiente* sobre la(s) *covariable(s)* a fin de obtener el *coeficiente de determinación* que expresa la proporción de varianza explicada por la(s) covariable(s). La diferencia existente entre esas dos regresiones, en términos de *coeficientes de determinación*, representa la proporción de varianza de la *variable dependiente* explicada por los grupos de *tratamiento* independientemente del influjo de la(s) *covariable(s)*. Por último, se lleva a cabo un análisis de la varianza sobre las puntuaciones a las que se ha eliminado la influencia de la(s) *covariable(s)* o posible(s) *variable(s) perturbadora(s)*.

No obstante, como señala López (1995), el procedimiento de ajuste para este tipo de diseños difiere en función de que los tratamientos administrados sean de naturaleza intersujetos o intrasujeto. En este sentido, cuando el análisis de la covarianza se aplica a diseños de medida múltiple, la covariable puede considerarse constante o variable, dependiendo de

la existencia de una sola medida objeto de control para cada sujeto o de una medida para la respuesta a cada uno de los tratamientos administrados. Esta consideración se refleja en modelos distintos para su análisis (Ato y López, 1992).

8.4.2. Diseño totalmente aleatorio, diseño de bloques aleatorios y diseño con covariables: breve análisis comparativo

Si se compara el *diseño con covariables* con el *diseño totalmente aleatorio*, se comprueba que el primero permite estimar los efectos experimentales con mayor precisión que el segundo. Al introducirse en el modelo ANCOVA un nuevo componente aditivo [$\beta(x_{ij} - \bar{x})$] (véase la Fórmula [8.67]), la varianza total ya no depende únicamente del error, ε_{ij} , y del parámetro α_j . De hecho, el componente aditivo añadido sustrae una determinada proporción de error, proporción que se incrementa a medida que aumenta la correlación que existe entre la covariable y la variable dependiente. En consecuencia, el diseño con covariables requiere un menor número de sujetos que el diseño totalmente aleatorio para comprobar la existencia de efectos significativos. Por otra parte, además de reducir el componente de error, el modelo de diseño experimental con covariables genera una estimación diferente de la magnitud del efecto experimental (Maxwell y Delaney, 1990). Ello se debe a que, mientras en el ANOVA que se aplica en el diseño totalmente aleatorio el efecto depende de la diferencia entre medias, en el ANCOVA se ajustan los datos a la recta de la regresión. Así, en este último caso, hay dos factores que afectan a las medias ajustadas: las diferencias entre las medias de la(s) covariable(s) y la pendiente de la recta de la regresión. Cuando las medias de la(s) covariable(s) son equivalentes en los distintos grupos, tanto el ANOVA como el ANCOVA producen la misma estimación del tamaño del efecto. Sin embargo, cuando los promedios son diferentes, con el ANCOVA se obtienen tamaños del efecto más pequeños que con el ANOVA.

Por otra parte, cuando se comparan entre sí el *diseño con covariables* y el *diseño de bloques aleatorios*, la mayoría de los autores se muestran de acuerdo en afirmar, que la principal ventaja que presenta la técnica de bloqueo con respecto al ANCOVA radica en que dicha técnica no requiere hacer suposiciones acerca de la forma que adopta la recta de la regresión de las puntuaciones de la variable dependiente sobre las puntuaciones de la(s) covariable(s) (Feldt, 1958; Huitema, 1980). Centrándose en criterios operativos y de carácter estadístico que pueden resultar útiles para seleccionar una u otra estrategia de análisis, Cox (1957) y Feldt (1958) afirman que el ANCOVA es más potente que el bloqueo, cuando la correlación existente entre la covariable y la variable dependiente es mayor que 0,80. Por el contrario, con correlaciones entre X e Y menores que 0,40, los dos autores recomiendan la técnica de bloqueo. Arnau (1986) establece la comparación entre estos dos modelos de diseño partiendo del tipo de relación existente entre X e Y . Así, el autor considera que si la relación entre X e Y es lineal, ambos diseños poseen la misma eficacia. No obstante, si la relación no es de naturaleza lineal, el diseño de bloques permite explicar una mayor cantidad de varianza que el diseño con covariables.

8.4.3. El análisis de datos en los diseños con covariables: análisis de la covarianza (ANCOVA)

8.4.3.1. Supuestos básicos del análisis de la covarianza

Al igual que el ANOVA, el ANCOVA es un procedimiento de análisis robusto a la violación de la normalidad y de la homogeneidad de las varianzas, cuando se trabaja con diseños

equilibrados. Sin embargo, no es una técnica robusta frente al incumplimiento del supuesto de independencia. Además, la correcta aplicación de esta estrategia de análisis requiere el cumplimiento de una serie de supuestos específicos, adicionales a los requeridos en el caso del ANOVA. A continuación abordamos cada uno de tales supuestos.

1. Supuesto de linealidad entre la variable dependiente y la covariable

La relación entre la variable dependiente y la covariable debe ser estrictamente lineal ya que, como señalan Huitema (1980) y Maxwell y Delaney (1990), cuando la relación no es lineal, el ajuste sobre la media es impropio. En consecuencia, si no se cumple dicho supuesto, los datos deben transformarse o debe recurrirse a la utilización de otro tipo de modelo de análisis estadístico.

2. Supuesto de independencia entre la variable independiente y la covariable

El cumplimiento de este supuesto implica que existe ortogonalidad entre la variable independiente y la covariable, a saber, que los tratamientos no ejercen ningún efecto sobre la covariable. En caso de existir dependencia entre la variable independiente y la covariable, sería cuestionable la utilización del ANCOVA (Huitema, 1980; Keppel, 1982).

Huitema (1980) expone de forma clara las consecuencias que puede implicar el incumplimiento de esta condición. En concreto, argumentan que en caso de que el tratamiento experimental ejerza influencia tanto sobre la variable dependiente como sobre la covariable, al ajustar estadísticamente la variable dependiente a partir de la covariable, se sustrae una parte del efecto experimental generado por el tratamiento y, por tanto, se produce un sesgo en la estimación de dicho efecto. Para evitar la confusión entre el efecto del tratamiento y los cambios asociados a la covariable, es necesario medir esta última antes de administrar el tratamiento e, incluso cumpliéndose dicho requisito, resulta muy aconsejable comprobar el supuesto de independencia entre la variable independiente y la covariable. Para ello, basta con considerar a la covariable (que es una variable de naturaleza continua) como variable dependiente y comprobar, mediante un ANOVA, que no existen diferencias estadísticamente significativas entre los promedios de la covariable en los distintos grupos de tratamiento. Es decir, la aceptación de la hipótesis de nulidad en el ANOVA permite asumir que existe independencia estadística entre la variable independiente y la covariable o, lo que es lo mismo, que los distintos grupos de tratamiento son equivalentes con respecto a la covariable.

El siguiente ejemplo puede ayudarnos a comprender en qué consiste este supuesto y a vislumbrar, con mayor claridad, la incidencia que puede tener su incumplimiento en la interpretación de los resultados. Supongamos que realizamos una investigación para estudiar los efectos que produce el tipo de examen (variable independiente) en el rendimiento de los alumnos (variable dependiente). Para ello, asignamos aleatoriamente la mitad de los estudiantes al grupo que debe realizar un examen tipo test y, la otra mitad, al grupo que debe realizar una prueba de preguntas abiertas. Es evidente que los sujetos acceden al examen con diferentes niveles de ansiedad (covariable), lo cual suponemos que puede afectar a su rendimiento. Si se cumple el supuesto que estamos abordando, a saber, si el tipo de examen (variable independiente) no afecta al nivel de ansiedad de los sujetos (covariable), cabe esperar que el ANCOVA, una vez controlada la influencia de la covariable, estime, de forma insesgada, el efecto del tipo de examen en el rendimiento de los sujetos. Por el contrario, ante el incumplimiento de este supuesto, es decir, si el tipo de examen genera distintos niveles de ansiedad (por ejemplo, si el hecho de someterse a una prueba tipo test produce mayor

ansiedad que el hecho de responder a una serie de preguntas abiertas), sería cuestionable utilizar el nivel de ansiedad (medido después de realizar la prueba) como covariable, ya que una parte de la variabilidad observada en el rendimiento (variable dependiente) que se debe a la ansiedad (covariable), está causada por el tipo de examen (variable independiente) que realiza el sujeto. Dicho de otro modo, la utilización del ANCOVA produciría un sesgo en los resultados, ya que una parte de la variabilidad que controlamos (varianza sistemática secundaria) constituiría una parte de la varianza sistemática primaria o de la variabilidad debida al efecto del tratamiento.

3. Supuesto de homogeneidad de las pendientes de regresión

Las pendientes de regresión (β_j') deben ser iguales para cada grupo de tratamiento. Si las líneas de regresión de los grupos no son paralelas, significa que existe un efecto de interacción entre los tratamientos y la covariable, es decir, que el efecto de la variable manipulada depende de los valores que obtiene cada sujeto en la covariable, lo que puede dificultar en gran medida la interpretación de los resultados. Como cabe deducir de lo que acabamos de afirmar, la consecuencia derivada de la violación de este supuesto no es de naturaleza técnica, sino teórica. De hecho, su incumplimiento no invalida necesariamente el proceso de estimación y de comprobación de hipótesis. Sin embargo, implica un cambio importante con respecto a la función de la covariable dentro del diseño ya que, en lugar de ser una variable extraña susceptible de control, se convierte en una variable explicativa adicional. En consecuencia, si no se cumple este supuesto, resulta esencial interpretar los efectos de interacción, dejando los efectos principales en un segundo plano.

La comprobación de este supuesto requiere verificar que la relación entre la covariable y la variable dependiente se mantiene constante en los distintos grupos experimentales. Así, se deben calcular las varianzas compartidas entre ambas, dentro de las diferentes condiciones de tratamiento y comprobar si tales varianzas son equivalentes. En caso afirmativo, se cumple el supuesto de homogeneidad de las pendientes de regresión. Si las varianzas compartidas entre la covariable y la variable dependiente son idénticas en todos los tratamientos, entonces las pendientes de regresión dentro de los diferentes grupos (intragrupo) serán idénticas y, debido a este hecho, el presente supuesto también se conoce como *prueba de homogeneidad de las pendientes intragrupo*.

Como señalan Pascual, Frías y García (1996), en el caso de que no se cumpla tal supuesto, no puede aplicarse el análisis de la covarianza clásico. De hecho, algunos autores consideran que, el supuesto de homogeneidad de las pendientes de regresión es un supuesto esencial para poder aplicar correctamente el ANCOVA (Cohen y Cohen, 1975; Kirk, 1982). No obstante, cabe señalar que Rogosa (1980) ha propuesto un método para desarrollar dicho análisis ante el incumplimiento del principio de homogeneidad de las pendientes intragrupo.

4. Supuesto de fiabilidad en la medida de la covariable

Este supuesto hace referencia a la necesidad de que la covariable sea medida sin error, es decir, al hecho de que exista fiabilidad en su medida. La transgresión de este supuesto, en los diseños no aleatorizados, produce sesgos en la estimación de las pendientes de regresión de los grupos, aumentando la probabilidad de cometer un error de tipo I. Por su parte, en los diseños experimentales, el incumplimiento de este supuesto produce una pérdida de potencia en la estimación de los efectos del tratamiento (Huitema, 1980).

Pascual, García y Frías (1995), entre otros, proporcionan varios ejemplos en los que desarrollan de manera muy didáctica algunos de los procedimientos que se pueden utilizar para comprobar si los datos cumplen los requisitos necesarios para la correcta aplicación del análisis de la covarianza.

8.4.3.2. Modelo general de análisis

El modelo analítico que se utiliza habitualmente para llevar a cabo la **prueba de la hipótesis** en los diseños con covariables es el *análisis de la covarianza*. La ecuación estructural asociada a la predicción que se realiza bajo la hipótesis alternativa en el *diseño de covarianza unifactorial intersujetos* adopta la siguiente expresión:

$$y_{ij} = \mu + \alpha_j + \beta(x_{ij} - \bar{x}) + \varepsilon_{ij} \quad (8.67)$$

donde:

y_{ij} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable de tratamiento.

μ = Media general de la variable dependiente.

α_j = Efecto principal asociado a la administración del j -ésimo nivel de la variable de tratamiento.

β = Coeficiente de regresión de la variable dependiente sobre la covariable.

x_{ij} = Puntuación obtenida en la covariable por el sujeto i bajo el j -ésimo tratamiento.

$\bar{x}$ = Media de las puntuaciones obtenidas por los sujetos en la covariable.

ε_{ij} = Término de error o componente aleatorio del modelo.

Si trabajamos con un *diseño intrasujeto*, la estructura más simple de un modelo con covariables incorpora una variable de naturaleza cuantitativa (X), un factor experimental cuyos niveles son aplicados a los mismos sujetos (A) y un factor sujeto (S). Como señala López (1995), para ejercer control estadístico sobre los tratamientos administrados, la covariable debe ser diferente para cada medida repetida (*modelo con covariable variable*), ya que si la covariable fuese constante para todas las medidas repetidas (*modelo con covariable constante*), no ejercería influencia sobre las fuentes de variación intrasujeto. La característica esencial del *modelo con covariable variable* es la inclusión de un efecto debido a la covariable de carácter intersujetos y otro de carácter intrasujeto. El primero de ellos hace referencia a la corrección que es necesario practicar sobre la suma de cuadrados debida a los sujetos y el segundo a la corrección requerida en el factor intrasujeto, corrección que supone la aplicación de un control estadístico distinto en cada uno de sus niveles. Así, la ecuación estructural asociada a la predicción que se realiza bajo la hipótesis alternativa en este tipo de modelo responde a la siguiente expresión:

$$y_{ij} = \mu + \alpha_j + \beta_{\text{inter}}(\bar{x}_j - \bar{x}) + \eta_i + \beta_{\text{intra}}(x_{ij} - \bar{x}_j) + \alpha\eta_{ji} + \varepsilon_{ij} \quad (8.68)$$

donde los componentes adicionales:

η_i = Efecto principal debido al i -ésimo sujeto.

β_{inter} = Coeficiente de regresión para las fuentes de variación intersujetos.

β_{intra} = Coeficiente de regresión para las fuentes de variación intrasujeto.

$\alpha\eta_{ji}$ = Interacción entre el factor de tratamiento y el factor sujeto.

ε_{ij} = Término de error del modelo.

En este caso, el ajuste del modelo requiere estimar dos coeficientes de regresión para una misma variable. López (1995) afirma que dicho problema se puede resolver planteando dos modelos distintos, de manera que el error de uno de ellos corresponda a las fuentes de variación intersujetos y el error del otro a las fuentes de variación intrasujeto.

Aunque no vamos a formular el modelo matemático subyacente al *diseño con covariables mixto*, es importante señalar que, en este tipo de diseños, las covariables pueden ser comunes o distintas para cada medida repetida. En el caso de que sean comunes, se consigue un control estadístico sobre las fuentes de variación intersujetos, mientras que si su naturaleza es variable, se controlan tanto las fuentes intersujetos como las intrasujeto (López, 1995).

Existen diversos *algoritmos de cálculo para desarrollar el análisis de la covarianza*. Arnau (1986), por ejemplo, tomando como referencia un diseño unifactorial intersujetos con una única covariable, comienza calculando la suma de cuadrados total, la intergrupos y la intragrupos correspondientes a la covariable (X) y a la variable dependiente (Y) y las sumas de productos cruzados entre X e Y . A partir de tales cálculos, estima las sumas cuadráticas de la variable dependiente ajustadas al efecto de la covariable. A continuación, obtiene las medias cuadráticas asociadas al componente sistemático y al componente aleatorio de la ecuación estructural, dividiendo las sumas de cuadrados ajustadas entre sus correspondientes grados de libertad. Por último, divide la MC_{entre} ajustada entre la MC_{intra} ajustada para calcular la razón F del análisis de la covarianza y llevar a cabo la prueba de la hipótesis.

Pascual (1995b), por su parte, desarrolla el ANCOVA desde la perspectiva de la comparación de modelos descritos en términos de efectos. El autor sigue la estrategia propuesta por Maxwell y Delaney (1990) que, en términos generales, incluye las siguientes etapas:

- Estimar el intercepto (α) y la pendiente (β) de la recta de regresión para la muestra total de los sujetos.
- Estimar el valor predicho por la ecuación de regresión para cada sujeto (y'_{ij}), así como la diferencia entre la puntuación real obtenida por el sujeto (y_{ij}) y tal valor, es decir, el error de predicción ($y_{ij} - y'_{ij}$). Estos cálculos se realizan bajo el supuesto de que todos los sujetos pertenecen al mismo grupo o, en otras palabras, bajo el supuesto de la hipótesis nula.
- Estimar la pendiente (β_j) de la recta de regresión para cada grupo de tratamiento, la pendiente media de regresión intragrupos y el intercepto (α_j) para cada grupo.
- Calcular la puntuación predicha (y'_{ij}) y el error de predicción ($y_{ij} - y'_{ij}$) correspondientes a cada sujeto en cada uno de los grupos de tratamiento. Estos errores de estimación son los errores que se cometen bajo el supuesto de la hipótesis alternativa.
- Ponderar las cantidades de error predichas bajo el supuesto de la hipótesis nula y bajo el supuesto de la hipótesis alternativa por sus correspondientes grados de libertad y obtener la razón F para llevar a cabo la prueba de la hipótesis.

En el siguiente epígrafe desarrollamos el ANCOVA en un diseño unifactorial intersujetos con una sola covariable, mediante los dos algoritmos de cálculo arriba descritos.

8.4.3.3. Ejemplo práctico

Supongamos que, en el ámbito de la psicología educativa, realizamos una investigación para examinar la influencia que ejerce la *hora en la que se imparte la docencia* (factor A) sobre el *rendimiento presentado por un grupo de niños en Matemáticas* (variable dependiente Y). En concreto, se desea examinar si dicho rendimiento varía en función de que las clases

se reciban por la mañana o por la tarde. No obstante, se considera que una variable extraña capaz de contaminar los resultados del estudio es el *nivel de rendimiento alcanzado por cada uno de los niños, en Matemáticas, antes de iniciar la investigación* (covariable o factor X). Con objeto de controlar tal variable, se decide aplicar el análisis de la covarianza a los datos obtenidos en el estudio. Para ello, se registran las calificaciones obtenidas por diez niños en una prueba que evalúa su nivel de conocimiento en Matemáticas, antes de aplicar los tratamientos experimentales. Posteriormente, se divide la muestra en dos grupos de cinco sujetos cada uno, recibiendo el primero de tales grupos la *docencia por la mañana* (a_1) y el segundo *por la tarde* (a_2), durante seis meses. Tras dicho período temporal, se vuelve a registrar el nivel de conocimiento que presentan los niños en Matemáticas. En la Tabla 8.22 pueden observarse los resultados obtenidos por los sujetos, tanto en el pretest o en la covariable X , como en el posttest o en la variable dependiente Y .

TABLA 8.22 Matriz de datos del experimento

A (Horario en el que se imparte la docencia)			
a_1 (Por la mañana)		a_2 (Por la tarde)	
Pretest X	Postest Y	Pretest X	Postest Y
7	9	8	12
6	10	7	8
4	7	6	7
8	9	9	11
5	8	7	10
$\Sigma X = 30$	$\Sigma Y = 43$	$\Sigma X = 37$	$\Sigma Y = 48$
$\Sigma XY = 263$		$\Sigma XY = 363$	
$\Sigma X \Sigma Y = 1.290$		$\Sigma X \Sigma Y = 1.776$	

La Tabla 8.22 refleja la estructura que corresponde a este modelo de diseño. En las filas se representan las *calificaciones obtenidas por los sujetos en Matemáticas*. Hemos de tener en cuenta que, independientemente del tratamiento experimental que reciba, a saber, a_1 o a_2 , a cada uno de los niños le corresponden dos puntuaciones: X e Y . Antes de abordar el ANCOVA y la verificación de los supuestos necesarios para su aplicación, en la Tabla 8.23 presentamos la estructura general de los datos correspondientes a un diseño de covarianza unifactorial con una sola covariable. Como ya es sabido, el subíndice $i = 1, 2, \dots, n$ corresponde a los sujetos y el subíndice $j = 1, 2, \dots, a$ hace referencia a los tratamientos.

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño, procederemos a desarrollar el ANCOVA, aplicando los algoritmos de cálculo propuestos por Arnau (1986) y por Pascual (1995b) y Maxwell y Delaney (1990), que ya han sido descritos en el subapartado correspondiente al «modelo general de análisis» dentro de este mismo punto (Epígrafe 8.4.3.2). No obstante, antes de aplicar el ANCOVA a los datos del ejemplo práctico que acabamos de plantear, examinaremos si tales datos cumplen con los

TABLA 8.23 Datos correspondiente a un diseño de covarianza unifactorial con una sola covariable: modelo general

A (Variable independiente)							
a_1		...	a_j		...	a_a	
X	Y		X	Y		X	Y
X_{11}	Y_{11}		X_{1j}	Y_{1j}		X_{1a}	Y_{1a}
...	...		...	...		...	...
X_{i1}	Y_{i1}		X_{ij}	Y_{ij}		X_{ia}	Y_{ia}
...	...		...	...		...	...
X_{n1}	Y_{n1}		X_{nj}	Y_{nj}		X_{na}	Y_{na}
ΣX	ΣY		ΣX	ΣY		ΣX	ΣY
ΣXY			ΣXY			ΣXY	
$\Sigma X \Sigma Y$			$\Sigma X \Sigma Y$			$\Sigma X \Sigma Y$	

requisitos necesarios para poder desarrollar dicho procedimiento de análisis. Recordemos que los supuestos básicos del ANCOVA son cuatro, a saber, el supuesto de linealidad entre la variable dependiente y la covariable, el supuesto de independencia entre la variable independiente y la covariable, el supuesto de homogeneidad de las pendientes de regresión y el supuesto de fiabilidad en la medida de la covariable. A lo largo de las siguientes páginas ilustramos los pasos necesarios para verificar tales supuestos.

• **Verificación de los supuestos del ANCOVA**

Tomando como referencia la matriz de datos de la Tabla 8.22 calculamos, en primer lugar, los errores correspondientes, tanto al modelo de predicción bajo la hipótesis nula, como bajo la hipótesis alternativa².

$$E_{H_0} = \{\varepsilon_{H_0}\} = \begin{matrix} & Y & X & \bar{y} & \bar{x} \\ \begin{bmatrix} 9 & 7 \\ 10 & 6 \\ 7 & 4 \\ 9 & 8 \\ 8 & 5 \\ 12 & 8 \\ 8 & 7 \\ 7 & 6 \\ 11 & 9 \\ 10 & 7 \end{bmatrix} & - & \begin{bmatrix} 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \end{bmatrix} & = & \begin{bmatrix} -0,1 & 0,3 \\ 0,9 & -0,7 \\ -2,1 & -2,7 \\ -0,1 & 1,3 \\ -1,1 & -1,7 \\ 2,9 & 1,3 \\ -1,1 & 0,3 \\ -2,1 & -0,7 \\ 1,9 & 2,3 \\ 0,9 & 0,3 \end{bmatrix} \end{matrix} \tag{8.69}$$

² El lector interesado en acceder al cálculo de los vectores de los errores de estimación correspondientes a los modelos de predicción bajo las hipótesis nula y alternativa, puede consultar el subapartado «Procedimiento 3: Desarrollo mediante vectores» del Epígrafe 6.1.2.2 del Capítulo 6.

$$E_{H_1} = \{\varepsilon_{H_1}\} = \begin{matrix} Y & X & \bar{y} & \bar{x} & (\bar{y}_j - \bar{y})(\bar{x}_j - \bar{x}) \\ \begin{bmatrix} 9 & 7 \\ 10 & 6 \\ 7 & 4 \\ 9 & 8 \\ 8 & 5 \\ 12 & 8 \\ 8 & 7 \\ 7 & 6 \\ 11 & 9 \\ 10 & 7 \end{bmatrix} & - & \begin{bmatrix} 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \\ 9,1 & 6,7 \end{bmatrix} & - & \begin{bmatrix} -0,5 & -0,7 \\ -0,5 & -0,7 \\ -0,5 & -0,7 \\ -0,5 & -0,7 \\ -0,5 & -0,7 \\ 0,5 & 0,7 \\ 0,5 & 0,7 \\ 0,5 & 0,7 \\ 0,5 & 0,7 \\ 0,5 & 0,7 \end{bmatrix} & = & \begin{bmatrix} 0,4 & 1 \\ 1,4 & 0 \\ -1,6 & -2 \\ 0,4 & 2 \\ -0,6 & -1 \\ 2,4 & 0,6 \\ -1,6 & -0,4 \\ -2,6 & -1,4 \\ 1,4 & 1,6 \\ 0,4 & -0,4 \end{bmatrix} \end{matrix} \quad (8.70)$$

A continuación calcularemos la SC_Y , la SC_X y la suma de productos cruzados entre X e Y , SP_{XY} , necesarias para el cálculo de la suma cuadrática total (modelo de la hipótesis nula) y de la suma cuadrática del error (modelo de la hipótesis alternativa). El valor de la SC_Y se obtiene elevando al cuadrado cada una de las puntuaciones de error de la columna correspondiente al factor Y y sumando entre sí los valores obtenidos. La SC_X se obtiene mediante el mismo procedimiento, pero aplicado a las puntuaciones de error de la columna correspondiente a la covariable X . Por último, calcularemos la suma de productos cruzados SP_{XY} , sumando entre sí las puntuaciones obtenidas tras multiplicar cada error correspondiente a la columna Y por su correlativo de la columna X .

Por tanto:

$$SC_T = E_{H_0}^2 = \begin{matrix} SC_Y & SP_{XY} \\ \begin{pmatrix} 24,9 & 16,3 \\ 16,3 & 20,1 \end{pmatrix} & \\ SP_{XY} & SC_X \end{matrix} \quad (8.71)$$

$$SC_e = E_{H_1}^2 = \begin{matrix} SC_Y & SP_{XY} \\ \begin{pmatrix} 22,4 & 12,8 \\ 12,8 & 15,2 \end{pmatrix} & \\ SP_{XY} & SC_X \end{matrix} \quad (8.72)$$

Finalmente, calcularemos el valor de la SC_{entre} siguiendo el mismo procedimiento, teniendo en cuenta que las puntuaciones necesarias para dicho cálculo se obtienen, en el caso de la columna correspondiente a la variable Y , mediante la fórmula $(\bar{y}_j - \bar{y})$, y en el caso de la columna correspondiente a la covariable X , mediante la fórmula $(\bar{x}_j - \bar{x})$ (esta matriz de puntuaciones ya ha sido calculada en el cómputo de E_{H_1} (8.70)). La matriz correspondiente a la SC_{entre} también puede obtenerse, de forma más sencilla, restando entre sí las matrices correspondientes a la SC_T y a la SC_e .

A continuación, elaboramos la tabla resumen de la descomposición de la suma de cuadrados de la variable dependiente en función de estos dos modelos y ajustando tales sumas de cuadrados al efecto de la covariable.

TABLA 8.24 Descomposición de la suma de cuadrados de la variable dependiente en función del modelo de predicción bajo la hipótesis nula y bajo la hipótesis alternativa y ajustando tales sumas de cuadrados al efecto de la covariable

Sumas de cuadrados	r_{XY}	η_{XY}^2	SC_Y	SC_{YX}	SC_Y^*
$SC_T = \begin{pmatrix} 24,9 & 16,3 \\ 16,3 & 20,1 \end{pmatrix}$	0,728	0,53	24,9 gl = $an - 1 = 9$	13,22 gl = $C = 1$	11,682 gl = $an - 1 - C = 8$
$SC_e = \begin{pmatrix} 22,4 & 12,8 \\ 12,8 & 15,2 \end{pmatrix}$	0,694	0,481	22,4 gl = $a(n - 1) = 8$	10,779 gl = $C = 1$	11,62 gl = $a(n - 1) - C = 7$
$SC_{\text{entre}} = \begin{pmatrix} 2,5 & 3,5 \\ 3,5 & 4,9 \end{pmatrix}$	—	—	2,5 gl = $a - 1 = 1$	—	0,06

Donde:

- $r_{XY} = \frac{SP_{XY}}{\sqrt{SC_Y SC_X}}$, siendo SP_{XY} la suma de productos cruzados entre X e Y, SC_Y la suma cuadrática de Y y SC_X la suma cuadrática de X. El índice r_{XY} hace referencia a la *correlación existente entre la variable dependiente y la covariable*.
- $\eta_{XY}^2 = r_{XY}^2$, que expresa el *porcentaje de varianza compartida entre la variable dependiente y la covariable*, es decir, la proporción de varianza de la variable dependiente explicada por la covariable.
- SC_Y = *Suma cuadrática de la variable dependiente*.
- $SC_{YX} = SC_Y \cdot \eta_{XY}^2$, que representa la *suma de cuadrados de la variable dependiente que se halla determinada por la relación que ésta mantiene con la covariable*.
- $SC_Y^* = SC_Y \cdot \Lambda_{XY}$, siendo $\Lambda_{XY} = 1 - \eta_{XY}^2$. El índice SC_Y^* hace referencia a la *suma de cuadrados de la variable dependiente ajustada al efecto de la covariable*.

Como ocurre en el caso de las sumas cuadráticas no ajustadas, en la descomposición correspondiente a las sumas de cuadrados ajustadas también se cumple la siguiente igualdad:

$$SC_{\text{entre}}^* = SC_T^* - SC_e^* \Rightarrow 11,682 - 11,62 = 0,06 \quad (8.73)$$

A partir de los índices que hemos calculado para elaborar la Tabla 8.24, examinaremos si nuestros datos cumplen con los requisitos necesarios para aplicar el ANCOVA. No obstante, antes de entrar en los supuestos propiamente dichos, resulta conveniente comprobar si la relación existente entre la variable dependiente y la covariable es estadísticamente significativa ya que, a medida que aumenta dicha relación, mayor es la reducción que se obtiene en el componente de varianza residual del modelo, al aplicar el ANCOVA. Para realizar dicha comprobación, aplicamos un análisis de la varianza a fin de dilucidar si la suma cuadrática residual de la variable dependiente, que se halla determinada por su relación con la covariable (SC_{eYX}), es significativamente distinta de cero (véase la Tabla 8.25).

Dado que el valor crítico que delimita la región de rechazo, $F_{\text{crít.}(0,05; 1,7)} = 5,59$, es menor que el valor observado, $F_{\text{obs.}} = 6,49$, debemos rechazar el modelo de la hipótesis nula como modelo explicativo de la relación existente entre la variable dependiente y la covariable. Por tanto, cabe concluir que existe una relación estadísticamente significativa entre ambas.

TABLA 8.25 Análisis de la varianza del modelo de regresión de Y sobre X , bajo el supuesto de la hipótesis alternativa

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	Razón F	p
X	10,779	$C = 1$	10,779	6,49	$< 0,05$
Error*	11,62	$a(n - 1) - C = 7$	1,66		
Error	22,4	$a(n - 1) = 8$			

- *Supuesto de linealidad entre la variable dependiente y la covariable*

La estrategia más sencilla para la comprobación de este supuesto consiste en la inspección visual de la representación gráfica del tipo de relación existente entre la variable dependiente y la covariable en cada uno de los grupos, partiendo de las puntuaciones directas obtenidas por los sujetos en tales grupos. Otra estrategia para examinar la linealidad consiste en representar gráficamente los valores residuales obtenidos a partir de la regresión de la variable dependiente sobre la covariable (Huitema, 1980). Una descripción más detallada acerca de los diversos procedimientos disponibles para la comprobación del supuesto de linealidad puede encontrarse en Draper y Smith (1981).

- *Supuesto de independencia entre la variable independiente y la covariable*

En el caso que nos ocupa, la finalidad del presente supuesto consiste en comprobar si las calificaciones obtenidas inicialmente por los sujetos en Matemáticas (puntuaciones en el pretest o en la covariable X) son equivalentes en los dos grupos establecidos en función del horario en el que se imparte la docencia (variable independiente o factor A). Para ello, realizamos el análisis de la varianza tomando, como variable independiente, la hora en la que se imparte la docencia y, como variable dependiente, las puntuaciones obtenidas por los sujetos en el pretest, es decir, en la covariable X .

TABLA 8.26 Análisis de la varianza tomando como variable independiente la hora en la que se imparte la docencia (factor A) y como variable dependiente las puntuaciones obtenidas por los sujetos en el pretest (covariable X)

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	Razón F	p
Intertratamientos (factor A)	$SCA = 4,9$	$a - 1 = 1$	$MCA = 4,9$	$F = 2,579$	$> 0,05$
Error o residual	$SCR = 15,2$	$a(n - 1) = 8$	$MCR = 1,9$		
Total	$SCT = 20,1$	$an - 1 = 9$			

Dado que el valor crítico que delimita la región de rechazo, $F_{\text{crít.}(0,05; 1,8)} = 5,32$, es mayor que el valor observado, $F_{\text{obs.}} = 2,579$, debemos aceptar el modelo de la hipótesis nula como

modelo explicativo de la relación existente entre la hora en la que se imparte la docencia y las calificaciones obtenidas por los sujetos en el pretest. Por tanto, cabe concluir que se cumple el supuesto de igualdad entre las puntuaciones presentadas por los sujetos en la covariable, en las distintas condiciones experimentales.

• *Supuesto de homogeneidad de las pendientes de regresión*

En el caso que nos ocupa, el objetivo del presente supuesto consiste en comprobar si la relación existente entre las puntuaciones obtenidas por los sujetos en el pretest (covariable X) y en el posttest (variable dependiente Y) se mantiene constante en los dos grupos experimentales. En concreto, la proporción de varianza residual compartida entre X e Y , a saber, $\eta_{XY}^2 = 0,481$, tiene que ser equivalente, tanto cuando los niños reciben sus clases por la mañana, como cuando las reciben por la tarde. Dicha equivalencia permite concluir que el tratamiento no interactúa con la covariable. Calculemos la varianza compartida entre X e Y en el primer tratamiento.

$$SC_{e(a_1)} = \begin{pmatrix} 0,4 & 1,4 & -1,6 & 0,4 & -0,6 \\ 1 & 0 & -2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 0,4 & 1 \\ 1,4 & 0 \\ -1,6 & -2 \\ 0,4 & 2 \\ -0,6 & -1 \end{pmatrix} = \begin{pmatrix} 5,2 & 5 \\ 5 & 10 \end{pmatrix} \quad (8.74)$$

Por tanto:

$$r_{XY(a_1)} = \frac{SP_{XY}}{\sqrt{SC_Y SC_X}} = \frac{5}{\sqrt{(5,2) \cdot (10)}} = 0,693 \quad (8.75)$$

$$\eta_{XY(a_1)}^2 = 0,481 \quad (8.76)$$

La varianza compartida entre X e Y en el segundo tratamiento se calcula siguiendo la misma lógica que en el caso precedente.

$$SC_{e(a_2)} = \begin{pmatrix} 2,4 & -1,6 & -2,6 & 1,4 & 0,4 \\ 0,6 & -0,4 & -1,4 & 1,6 & -0,4 \end{pmatrix} \begin{pmatrix} 2,4 & 0,6 \\ -1,6 & -0,4 \\ -2,6 & -1,4 \\ 1,4 & 1,6 \\ 0,4 & -0,4 \end{pmatrix} = \begin{pmatrix} 17,2 & 7,8 \\ 7,8 & 5,2 \end{pmatrix} \quad (8.77)$$

Por tanto:

$$r_{XY(a_2)} = \frac{SP_{XY}}{\sqrt{SC_Y SC_X}} = \frac{7,8}{\sqrt{(17,2) \cdot (5,2)}} = 0,825 \quad (8.78)$$

$$\eta_{XY(a_2)}^2 = 0,68 \quad (8.79)$$

Tras realizar estos cálculos, elaboramos la tabla resumen de la descomposición de la suma de cuadrados del error, en las dos condiciones experimentales.

TABLA 8.27 Descomposición de la suma de cuadrados del error en las dos condiciones experimentales

Sumas de cuadrados	r_{XY}	η_{XY}^2	SC_Y	SC_{YX}	SC_Y^*
$SC_e = \begin{pmatrix} 22,4 & 12,8 \\ 12,8 & 15,2 \end{pmatrix}$	0,694	0,481	22,4	10,779	11,62
$SC_{e(a_1)} = \begin{pmatrix} 5,2 & 5 \\ 5 & 10 \end{pmatrix}$	0,693	0,481	5,2	2,5	2,699
$SC_{e(a_2)} = \begin{pmatrix} 17,2 & 7,8 \\ 7,8 & 5,2 \end{pmatrix}$	0,825	0,680	17,2	11,696	5,504

Como cabe observar en la Tabla 8.27, la varianza residual compartida entre la variable dependiente y la covariable, en ambos grupos conjuntamente, es de 0,481. A su vez, en el primer grupo adopta un valor de 0,481 y, en el segundo grupo, un valor de 0,680.

Con respecto a las sumas de cuadrados ajustadas al efecto de X en cada tratamiento, cabe apreciar que la $SC_{e(a_1)}^*$ es igual a 2,699 y que la $SC_{e(a_2)}^*$ adopta un valor de 5,504. Como se ha señalado al abordar los índices que configuran la Tabla 8.24, tales valores pueden obtenerse mediante las siguientes expresiones

$$SC_{e(a_1)}^* = SC_{e(a_1)} \cdot \Lambda_{XY(a_1)} = (5,2) \cdot (0,519) = 2,699 \quad (8.80)$$

$$SC_{e(a_2)}^* = SC_{e(a_2)} \cdot \Lambda_{XY(a_2)} = (17,2) \cdot (0,32) = 5,504 \quad (8.81)$$

La *suma cuadrática residual intragrupo* ($SC_{e(a)}^*$) se obtiene sumando las sumas de cuadrados residuales ajustadas, correspondientes a cada tratamiento, a saber:

$$SC_{e(a)}^* = \Sigma SC_{e(a_j)}^* = 2,699 + 5,504 = 8,203 \quad (8.82)$$

A su vez, el *efecto de interacción entre la covariable y el tratamiento* ($SC_{X \times A}^*$) se calcula mediante la siguiente expresión:

$$SC_{X \times A}^* = SC_{eY}^* - SC_{e(a)}^* = 11,62 - 8,203 = 3,417 \quad (8.83)$$

Así, la *varianza residual* se descompone en las siguientes fuentes de variación:

$$SC_e = SC_{eYX} + SC_{X \times A}^* + SC_{e(a)}^* \quad (8.84)$$

donde:

- SC_{eYX} representa la *suma de cuadrados residual de la variable dependiente que se halla determinada por su relación con la covariable*.
- $SC_{X \times A}^*$ corresponde a la *suma cuadrática residual asociada a la interacción entre la covariable y la variable independiente*.
- $SC_{e(a)}^*$ expresa la *suma de cuadrados correspondiente al componente residual del modelo*.

Llegados a este punto, la comprobación del supuesto sólo requiere verificar si la interacción entre la covariable y el tratamiento es, o no, estadísticamente significativa. Para ello, aplicamos el análisis de la varianza sobre las fuentes de variación de la Fórmula (8.84).

TABLA 8.28 ANOVA tomando como fuentes de variación los diferentes términos correspondientes a la varianza residual

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	Razón F	p
X	10,779	$C = 1$	10,779	$F_X = \frac{MC_X}{MC_{e(a)}} = 7,885$	$< 0,05$
$X \times A$	3,417	$a - 1 = 1$	3,417	$F_{X \times A} = \frac{MC_{X \times A}}{MC_{e(a)}} = 2,499$	$> 0,05$
Error (a)	8,203	$N - 2a = 6$	1,367		
Error	22,4	$N - a - 1 = 8$			

Dado que la F teórica correspondiente a la interacción entre la covariable y el tratamiento, $F_{\text{crit.}(0,05; 1,6)} = 5,98$, es mayor que la F observada, $F_{\text{obs.}} = 2,499$, cabe concluir que no existe efecto de interacción entre la variable independiente y la covariable, es decir, que la influencia que ejerce la hora en la que se imparte la docencia sobre la variable dependiente no se halla condicionada por las puntuaciones que obtienen los sujetos en el pretest. En consecuencia, podemos afirmar que se cumple el supuesto de homogeneidad de las pendientes de regresión.

• *Supuesto de fiabilidad en la medida de la covariable*

El problema del error o de la fiabilidad en la medida es una de las cuestiones centrales que aborda la *Psicometría*, disciplina a la que remitimos al lector interesado en la comprobación de este supuesto.

• **Análisis de la covarianza**

Procedimiento 1

En la Tabla 8.29 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas de la covariable X y de la variable dependiente Y , así como para el cálculo de las sumas de productos cruzados entre X e Y .

VARIABLE X O COVARIABLE

Comenzamos con el cálculo del término general C_X .

$$C_X = \frac{1}{an} \left(\sum_j^a \sum_i^n X_{ij} \right)^2 \tag{8.85}$$

$$C_X = \frac{1}{2 \cdot 5} (7 + 6 + 4 + 8 + 5 + 8 + 7 + 6 + 9 + 7)^2 = \frac{(67)^2}{10} = 448,9$$

TABLA 8.29 Fórmulas necesarias para el cálculo de las sumas cuadráticas de la covariable X y la variable dependiente Y , así como para el cálculo de las sumas de productos cruzados entre X e Y

Fuentes de variación	Sumas de cuadrados		Sumas de productos cruzados
	X	Y	XY
Factor A	$SCX = \frac{1}{n} \sum_j \left(\sum_i^n X_{ij} \right)^2 - C_X$	$SCY = \frac{1}{n} \sum_j \left(\sum_i^n Y_{ij} \right)^2 - C_Y$	$SCXY = \frac{1}{n} \sum_j \left(\sum_i^n X_{ij} \right) \left(\sum_i^n Y_{ij} \right) - C_{XY}$
Error	$SCR_X = SCT_X - SCX$	$SCR_Y = SCT_Y - SCY$	$SCR_{XY} = SCT_{XY} - SCXY$
Total	$SCT_X = \sum_j \sum_i^n X_{ij}^2 - C_X$	$SCT_Y = \sum_j \sum_i^n Y_{ij}^2 - C_Y$	$SCT_{XY} = \sum_j \sum_i^n X_{ij} Y_{ij} - C_{XY}$
Términos C	$C_X = \frac{1}{an} \left(\sum_j \sum_i^n X_{ij} \right)^2$	$C_Y = \frac{1}{an} \left(\sum_j \sum_i^n Y_{ij} \right)^2$	$C_{XY} = \frac{1}{an} \left(\sum_j \sum_i^n X_{ij} \right) \left(\sum_j \sum_i^n Y_{ij} \right)$

FACTOR A

Una vez estimado el término C_X , procedemos a calcular la variabilidad de la covariable explicada por la pertenencia a cada uno de los grupos establecidos en función del horario en el que se imparte la docencia.

$$SCX = \frac{1}{n} \sum_j \left(\sum_i^n X_{ij} \right)^2 - C_X \quad (8.86)$$

$$SCX = \frac{1}{5} [(7 + 6 + 4 + 8 + 5)^2 + (8 + 7 + 6 + 9 + 7)^2] - 448,9 = 453,8 - 448,9 = 4,9$$

Hacemos lo propio con la variabilidad total y con la variabilidad residual.

VARIABILIDAD TOTAL: SCT_X

$$SCT_X = \sum_j \sum_i^n X_{ij}^2 - C_X \quad (8.87)$$

$$SCT_X = [(7)^2 + (6)^2 + (4)^2 + (8)^2 + (5)^2 + (8)^2 + (7)^2 + (6)^2 + (9)^2 + (7)^2] - 448,9 = 469 - 448,9 = 20,1$$

ERROR

$$SCR_X = SCT_X - SCX = 20,1 - 4,9 = 15,2 \quad (8.88)$$

Llegados a este punto, hemos finalizado con los cálculos correspondientes a las sumas cuadráticas de la *covariable*.

VARIABLE Y O VARIABLE DEPENDIENTE

Como en el caso anterior, realizaremos los cálculos correspondientes a las sumas de cuadrados de la variable dependiente, comenzando con el término general C_Y .

$$C_Y = \frac{1}{an} \left(\sum_j^a \sum_i^n Y_{ij} \right)^2 \quad (8.89)$$

$$C_Y = \frac{1}{2 \cdot 5} (9 + 10 + 7 + 9 + 8 + 12 + 8 + 7 + 11 + 10)^2 = \frac{(91)^2}{10} = 828,1$$

A continuación, calculamos la variabilidad asociada al factor A, la variabilidad total y la variabilidad residual.

FACTOR A

$$SCY = \frac{1}{n} \sum_j^a \left(\sum_i^n Y_{ij} \right)^2 - C_Y \quad (8.90)$$

$$SCY = \frac{1}{5} [(9 + 10 + 7 + 9 + 8)^2 + (12 + 8 + 7 + 11 + 10)^2] - 828,1 = 830,6 - 828,1 = 2,5$$

VARIABILIDAD TOTAL: SCT_Y

$$SCT_Y = \sum_j^a \sum_i^n Y_{ij}^2 - C_Y \quad (8.91)$$

$$SCT_Y = [(9)^2 + (10)^2 + (7)^2 + (9)^2 + (8)^2 + (12)^2 + (8)^2 + (7)^2 + (11)^2 + (10)^2] - 828,1 = 853 - 828,1 = 24,9$$

ERROR

$$SCR_Y = SCT_Y - SCY = 24,9 - 2,5 = 22,4 \quad (8.92)$$

Tras el cálculo de las sumas cuadráticas de X y de Y, procedemos a estimar las sumas de productos cruzados entre X e Y.

SUMAS DE PRODUCTOS CRUZADOS**TÉRMINO C_{XY}**

$$C_{XY} = \frac{1}{an} \left(\sum_j^a \sum_i^n X_{ij} \right) \left(\sum_j^a \sum_i^n Y_{ij} \right) \quad (8.93)$$

$$C_{XY} = \frac{1}{2 \cdot 5} [7 + 6 + 4 + 8 + 5 + 8 + 7 + 6 + 9 + 7] \cdot (9 + 10 + 7 + 9 + 8 + 12 + 8 + 7 + 11 + 10) = \frac{(67) \cdot (91)}{10} = 609,7$$

FACTOR A

$$SCXY = \frac{1}{n} \sum_j^a \left(\sum_i^n X_{ij} \right) \left(\sum_i^n Y_{ij} \right) - C_{XY} \quad (8.94)$$

$$SCXY = \frac{1}{5} [(7 + 6 + 4 + 8 + 5)(9 + 10 + 7 + 9 + 8) + \\ + (8 + 7 + 6 + 9 + 7)(12 + 8 + 7 + 11 + 10)] - 609,7 = 613,2 - 609,7 = 3,5$$

VARIABILIDAD TOTAL: SCT_{XY}

$$SCT_{XY} = \sum_j^a \sum_i^n X_{ij} Y_{ij} - C_{XY} \quad (8.95)$$

$$SCT_{XY} = \left[(7) \cdot (9) + (6) \cdot (10) + (4) \cdot (7) + (8) \cdot (9) + (5) \cdot (8) + \right. \\ \left. + (8) \cdot (12) + (7) \cdot (8) + (6) \cdot (7) + (9) \cdot (11) + (7) \cdot (10) \right] - 609,7 = \\ = 626 - 609,7 = 16,3$$

ERROR

$$SCR_{XY} = SCT_{XY} - SCXY = 16,3 - 3,5 = 12,8 \quad (8.96)$$

Tras el cálculo de las *sumas de cuadrados* y de las *sumas de productos cruzados*, podemos estimar las *medias cuadráticas* o las *varianzas ajustadas al efecto de la covariable*. Para ello debemos conocer, en primer lugar, los *grados de libertad* correspondientes a cada uno de los *efectos*. Tras obtener dichos valores, podemos estimar la *razón entre varianzas* o el *estadístico F*.

Recordemos que el diseño que nos ocupa consta de dos fuentes de variación principales: el efecto del factor A y la variabilidad residual. La razón entre las varianzas ajustadas, correspondientes a tales fuentes de variación, permite saber si la variable manipulada ejerce una influencia estadísticamente significativa sobre la variable dependiente, tras someter a control estadístico el efecto de la covariable. A continuación, procedemos al cálculo de dichas varianzas.

MEDIAS CUADRÁTICAS O VARIANZAS AJUSTADAS AL EFECTO DE LA COVARIABLE

La *varianza ajustada del factor A*, $MCA_{Y'}$, se calcula mediante la siguiente expresión:

$$MCA_{Y'} = \frac{SCA_{Y'}}{a - 1} \quad (8.97)$$

donde:

$$SCA_{Y'} = SCT_Y - \frac{SCT_{XY}^2}{SCT_X} - SCR_Y + \frac{SCR_{XY}^2}{SCR_X} \quad (8.98)$$

La *varianza ajustada correspondiente al error*, $MCR_{Y'}$, se calcula aplicando la siguiente fórmula:

$$MCR_{Y'} = \frac{SCR_{Y'}}{a(n-1) - C} \quad (8.99)$$

donde:

$$SCR_{Y'} = SCR_Y - \frac{SCR_{XY}^2}{SCR_X} \quad (8.100)$$

En las Fórmulas (8.97) y (8.99), los *denominadores* representan los *grados de libertad* correspondientes a cada una de las fuentes de variación. Cabe observar, en las fórmulas precedentes, que para el cálculo de las *medias cuadráticas* o de las *varianzas ajustadas*, se toma en consideración la variabilidad asociada a la *covariable*. Antes de aplicar tales fórmulas a los datos de nuestro ejemplo, las recapitulamos en la Tabla 8.30.

TABLA 8.30 Análisis de la covarianza para un diseño de covarianza unifactorial intersujetos, con una sola covariable: modelo general

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A <i>ajustado</i>	$SCA_{Y'}$	$a - 1$	$MCA_{Y'} = \frac{SCA_{Y'}}{a - 1}$	$F_A = \frac{MCA_{Y'}}{MCR_{Y'}}$
Error <i>ajustado</i>	$SCR_{Y'}$	$a(n - 1) - C$	$MCR_{Y'} = \frac{SCR_{Y'}}{a(n - 1) - C}$	
Total	$SCT_{Y'}$	$(an) - 1 - C$		

Aplicando las Fórmulas (8.98) y (8.97) a los datos de nuestro ejemplo:

$$SCA_{Y'} = SCT_Y - \frac{SCT_{XY}^2}{SCT_X} - SCR_Y + \frac{SCR_{XY}^2}{SCR_X}$$

$$SCA_{Y'} = 24,9 - \frac{16,3^2}{20,1} - 22,4 + \frac{12,8^2}{15,2} = 0,06$$

Una vez conocida la suma cuadrática ajustada del factor A ($SCA_{Y'}$), podemos calcular la *varianza ajustada* de dicho factor ($MCA_{Y'}$).

$$MCA_{Y'} = \frac{SCA_{Y'}}{a - 1}$$

$$MCA_{Y'} = \frac{0,06}{2 - 1} = 0,06$$

Por otra parte, la suma cuadrática residual ajustada ($SCR_{Y'}$) y la varianza residual ajustada ($MCR_{Y'}$) se calculan aplicando las Fórmulas (8.100) y (8.99), respectivamente.

$$SCR_{Y'} = SCR_Y - \frac{SCR_{XY}^2}{SCR_X}$$

$$SCR_{Y'} = 22,4 - \frac{12,8^2}{15,2} = 11,62$$

A partir de dicho valor:

$$MCR_{Y'} = \frac{SCR_{Y'}}{a(n-1) - C}$$

$$MCR_{Y'} = \frac{11,62}{2(5-1) - 1} = 1,66$$

Llegados a este punto, cabe aplicar el ANCOVA a los datos de nuestro ejemplo práctico. En la Tabla 8.31 se presentan los términos necesarios para llevar a cabo la prueba de la hipótesis.

TABLA 8.31 Análisis de la covarianza para un diseño de covarianza unifactorial intersujetos, con una sola covariable: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A ajustado	$SCA_{Y'} = 0,06$	$a - 1 = 1$	$MCA_{Y'} = \frac{SCA_{Y'}}{a - 1} = 0,06$	$F_A = \frac{MCA_{Y'}}{MCR_{Y'}}$ $F_A = \frac{0,06}{1,66} = 0,036$
Error ajustado	$SCR_{Y'} = 11,62$	$a(n-1) - C = 7$	$MCR_{Y'} = \frac{SCR_{Y'}}{a(n-1) - C} = 1,66$	
Total	$SCT_{Y'} = 11,682$	$(an) - 1 - C = 8$		

Estableciendo un nivel de confianza del 95 % ($\alpha = 0,05$) y trabajando con una hipótesis de una sola cola, obtenemos el siguiente valor crítico del estadístico F , a saber, $F_{\text{crít.}(0,05; 1,7)} = 5,59$. Dado que el valor de la F observada, $F_{\text{obs.}} = 0,036$, es menor que el valor crítico, no podemos rechazar la hipótesis nula. En consecuencia, cabe concluir que, tras controlar estadísticamente el posible efecto debido al nivel de conocimiento que presentan los niños en Matemáticas antes de aplicar los tratamientos, el hecho de recibir la docencia por la mañana o por la tarde no ejerce una influencia estadísticamente significativa sobre la variable dependiente (rendimiento en Matemáticas).

Aunque suponemos que el lector ya se habrá percatado de ello, cabe señalar que las *sumas cuadráticas ajustadas* pueden obtenerse mediante procedimientos menos laboriosos

que el que acabamos de presentar. Uno de tales métodos ya ha sido expuesto al elaborar la Tabla 8.24 dentro de este mismo epígrafe (Epígrafe 8.4.3.3). Además de calcular las sumas de cuadrados ajustadas, correspondientes al factor A y al término de error, en la Tabla 8.24 también estimamos la variabilidad residual de la variable dependiente, que se halla determinada por su relación con la covariable, a saber, el término SC_{eYX} . Dicho término permite examinar la magnitud del componente de error, debido al efecto de la covariable. Incorporando tal componente a los elementos de la Tabla 8.31, obtenemos la Tabla 8.32.

TABLA 8.32 Análisis de la covarianza para un diseño unifactorial intersujetos con una sola covariable, añadiendo a las sumas cuadráticas ajustadas del factor A y del término residual, la variabilidad residual de la variable dependiente explicada por la covariable, bajo el supuesto de la hipótesis alternativa

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F	p
X	10,779	1	10,779	6,49	< 0,05
Factor A <i>ajustado</i>	0,06	1	0,06	0,036	> 0,05
Error <i>ajustado</i>	11,62	7	1,66		
Total	11,682	8			

Adviértase que la fuente de variación X correspondería a un ANCOVA distinto, donde las sumas cuadráticas del error ajustado y de la fuente de variación total adoptarían los valores de 11,62 y 22,4, respectivamente.

Como cabe apreciar en la Tabla 8.32, la covariable permite reducir, significativamente, el componente de varianza residual del modelo ($F_{\text{crít.}(0,05; 1,7)} = 5,59 < F_{\text{obs.}} = 6,49$). A su vez, tras ajustar las puntuaciones de la variable dependiente al efecto de la covariable, no se aprecian diferencias estadísticamente significativas en las calificaciones obtenidas en Matemáticas por los niños que reciben sus clases por la mañana, frente a los que las reciben por la tarde ($F_{\text{crít.}(0,05; 1,7)} = 5,59 > F_{\text{obs.}} = 0,036$).

Procedimiento 2

El procedimiento que se presenta para desarrollar el ANCOVA, en este subapartado, es totalmente diferente al que se ha expuesto previamente. En concreto, abordaremos el ANCOVA tomando como referencia el análisis de la regresión.

Como se ha señalado al plantear el modelo general de análisis para el diseño con covariables, el desarrollo del ANCOVA, mediante este procedimiento, incluye las siguientes etapas:

1. Estimar la pendiente (β) y el intercepto (α) de la recta de regresión para la muestra total de los sujetos. Tras realizar tales cálculos, se estima el valor predicho por la recta de regresión para cada sujeto (y'_{ij}), así como la diferencia entre la puntuación real obtenida por el sujeto (y_{ij}) y tal valor, es decir, el error de predicción ($y_{ij} - y'_{ij}$). Estos cálculos se realizan bajo el supuesto de que todos los sujetos pertenecen al mismo grupo o, en otras palabras, bajo el supuesto de la hipótesis nula.

2. Estimar la pendiente (β_j) de la recta de regresión para cada grupo de tratamiento, la pendiente media intragrupo y el intercepto (α_j) para cada grupo. A continuación, se calcula la puntuación predicha (y'_{ij}) y el error de predicción ($y_{ij} - y'_{ij}$) correspondientes a cada sujeto en cada uno de los grupos de tratamiento. Estos errores de estimación son los errores que se cometen bajo el supuesto de la hipótesis alternativa.
3. Ponderar las cantidades de error predichas, bajo el supuesto de la hipótesis nula y bajo el supuesto de la hipótesis alternativa, por sus correspondientes grados de libertad y obtener la razón F para llevar a cabo la prueba de la hipótesis.

Tras describir las etapas que se deben seguir para desarrollar el ANCOVA mediante este método, procedemos a abordarlas, tomando como referencia la matriz de datos correspondiente a la Tabla 8.22.

PRIMERA ETAPA

La pendiente (β) y el intercepto (α) de la recta de regresión se calculan mediante las siguientes fórmulas:

$$\beta = \sum_i^n \sum_j^a \frac{(X_{ij} - \bar{X})(Y_{ij} - \bar{Y})}{(X_{ij} - \bar{X})^2} \quad (8.101)$$

$$\alpha = \bar{Y} - \beta \bar{X} \quad (8.102)$$

En la Tabla 8.33 pueden observarse las operaciones necesarias para estimar tales valores.

TABLA 8.33 Cálculos basados en la regresión lineal para el desarrollo del ANCOVA (1): ejemplo práctico

X	Y	$(X_{ij} - \bar{X})$	$(Y_{ij} - \bar{Y})$	$(X_{ij} - \bar{X})^2$	$(X_{ij} - \bar{X})(Y_{ij} - \bar{Y})$
7	9	0,3	-0,1	0,09	-0,03
6	10	-0,7	0,9	0,49	-0,63
4	7	-2,7	-2,1	7,29	5,67
8	9	1,3	-0,1	1,69	-0,13
5	8	-1,7	-1,1	2,89	1,87
8	12	1,3	2,9	1,69	3,77
7	8	0,3	-1,1	0,09	-0,33
6	7	-0,7	-2,1	0,49	1,47
9	11	2,3	1,9	5,29	4,37
7	10	0,3	0,9	0,09	0,27
67	91	SUMATORIOS (Σ)		20,1	16,3

Suponemos que los lectores familiarizados con el análisis de la regresión asociarán fácilmente la Tabla 8.33 con el modelo de regresión lineal simple.

Partiendo de los sumatorios que se presentan en la última fila de la tabla, procedemos a calcular las medias aritméticas $\bar{X}$ e $\bar{Y}$.

$$\bar{X} = \frac{\sum_i^n X}{n} = \frac{67}{10} = 6,7 \quad \text{e} \quad \bar{Y} = \frac{\sum_i^n Y}{n} = \frac{91}{10} = 9,1$$

A continuación, estimamos los parámetros β y α de la recta de regresión, a partir de las Fórmulas (8.101) y (8.102).

$$\beta = \frac{\sum_i^n \sum_j^a \frac{(X_{ij} - \bar{X})(Y_{ij} - \bar{Y})}{(X_{ij} - \bar{X})^2}} = \frac{16,3}{20,1} = 0,811$$

$$\alpha = \bar{Y} - \beta\bar{X} = 9,1 - (0,811)(6,7) = 3,666$$

Ya sabemos que la *ecuación de la regresión lineal* viene expresada por $Y'_{ij} = \alpha + \beta X_{ij}$, donde cada puntuación Y'_{ij} hace referencia a la proporción de variabilidad de la puntuación real de cada sujeto, explicada por la covariable. A su vez, el *error de regresión* o la proporción de variabilidad no explicada por la covariable se expresa como $(\epsilon_0)_{ij} = Y_{ij} - Y'_{ij}$. Partiendo de tales expresiones, podemos elaborar la Tabla 8.34.

TABLA 8.34 Cálculos basados en la *regresión lineal* para el desarrollo del ANCOVA (2): ejemplo práctico

<i>X</i>	<i>Y</i>	$Y'_{ij} = \alpha + \beta X_{ij}$	$(\epsilon_0)_{ij} = Y_{ij} - Y'_{ij}$	$(\epsilon_0)_{ij}^2$
7	9	9,343	-0,343	0,117
6	10	8,532	1,468	2,155
4	7	6,91	0,09	0,008
8	9	10,154	-1,154	1,331
5	8	7,721	0,279	0,077
8	12	10,154	1,846	3,407
7	8	9,343	-1,343	1,803
6	7	8,532	-1,532	2,347
9	11	10,965	0,035	0,001
7	10	9,343	0,657	0,431
67	91	SUMATORIOS (Σ)	-0,003	11,6813

Por tanto:

$$\Sigma (\epsilon_0)_{ij}^2 = 11,6813$$

SEGUNDA ETAPA

Comenzaremos calculando la pendiente (β_j) en cada uno de los grupos de tratamiento. Para ello, aplicaremos la Fórmula (8.103) a los datos de nuestro ejemplo.

$$\beta_j = \sum_i^n \frac{(X_{ij} - \bar{X}_j)(Y_{ij} - \bar{Y}_j)}{(X_{ij} - \bar{X}_j)^2} \tag{8.103}$$

• **Primer grupo (*j* = 1)**

Elaboramos la tabla correspondiente (véase la Tabla 8.35).

TABLA 8.35 Cálculos basados en la *regresión lineal* para el desarrollo del ANCOVA (3): ejemplo práctico ($j = 1$)

X	Y	$(X_{i1} - \bar{X}_1)$	$(Y_{i1} - \bar{Y}_1)$	$(X_{i1} - \bar{X}_1)^2$	$(X_{i1} - \bar{X}_1)(Y_{i1} - \bar{Y}_1)$
7	9	1	0,1	1	0,4
6	10	0	1,4	0	0
4	7	-2	-1,6	4	3,2
8	9	2	0,4	4	0,8
5	8	-1	-0,6	1	0,6
30	43	SUMATORIOS (Σ)		10	5

Partiendo de los sumatorios que se presentan en la última fila de la tabla, procedemos a calcular las medias aritméticas $\bar{X}_1$ e $\bar{Y}_1$.

$$\bar{X}_1 = \frac{\sum_i^n X}{n} = \frac{30}{5} = 6 \quad \text{e} \quad \bar{Y}_1 = \frac{\sum_i^n Y}{n} = \frac{43}{5} = 8,6$$

Tras realizar dichos cálculos, estimamos la pendiente (β_j) de la recta de regresión en el primer grupo ($j = 1$), aplicando la Fórmula (8.104).

$$\beta_{j=1} = \sum_i^n \frac{(X_{i1} - \bar{X}_1)(Y_{i1} - \bar{Y}_1)}{(X_{i1} - \bar{X}_1)^2} \quad (8.104)$$

$$\beta_{j=1} = \frac{5}{10} = 0,5$$

• **Segundo grupo ($j = 2$)**

Para el desarrollo de los cálculos correspondientes a la segunda condición experimental, seguiremos los mismos pasos que en el caso del primer grupo de tratamiento.

TABLA 8.36 Cálculos basados en la *regresión lineal* para el desarrollo del ANCOVA (4): ejemplo práctico ($j = 2$)

X	Y	$(X_{i2} - \bar{X}_2)$	$(Y_{i2} - \bar{Y}_2)$	$(X_{i2} - \bar{X}_2)^2$	$(X_{i2} - \bar{X}_2)(Y_{i2} - \bar{Y}_2)$
8	12	0,6	2,4	0,36	1,44
7	8	-0,4	-1,6	0,16	0,64
6	7	-1,4	-2,6	1,96	3,64
9	11	1,6	1,4	2,56	2,24
7	10	-0,4	0,4	0,16	-0,16
37	48	SUMATORIOS (Σ)		5,2	7,8

Partiendo de los sumatorios que se presentan en la última fila de la tabla, procedemos a calcular las medias aritméticas $\bar{X}_2$ e $\bar{Y}_2$.

$$\bar{X}_2 = \frac{\sum_i^n X}{n} = \frac{37}{5} = 7,4 \quad \text{e} \quad \bar{Y}_2 = \frac{\sum_i^n Y}{n} = \frac{48}{5} = 9,6$$

Tras realizar dichos cálculos, estimamos la pendiente (β_j) en el segundo grupo ($j = 2$), aplicando la Fórmula (8.105).

$$\beta_{j=2} = \frac{\sum_i^n (X_{i2} - \bar{X}_2)(Y_{i2} - \bar{Y}_2)}{(X_{i2} - \bar{X}_2)^2} \quad (8.105)$$

$$\beta_{j=2} = \frac{7,8}{5,2} = 1,5$$

Una vez estimados los valores de la pendiente (β_j) para cada uno de los grupos experimentales, estimamos la pendiente media de regresión intragrupo (β_{intra}) aplicando la Fórmula (8.106).

$$\beta_{\text{intra}} = \frac{\sum (X_{i1} - \bar{X}_1)^2 \beta_1 + \sum (X_{i2} - \bar{X}_2)^2 \beta_2}{\sum (X_{i1} - \bar{X}_1)^2 + \sum (X_{i2} - \bar{X}_2)^2} \quad (8.106)$$

$$\beta_{\text{intra}} = \frac{(10)(0,5) + (5,2)(1,5)}{10 + 5,2} = 0,8421$$

A continuación, estimamos el intercepto (α_j) en cada uno de los grupos de tratamiento. Para ello, aplicamos la Fórmula (8.107) a los datos de nuestro ejemplo.

$$\alpha_j = \bar{Y}_j - \beta_{\text{intra}} \bar{X}_j \quad (8.107)$$

$$\alpha_1 = \bar{Y}_1 - \beta_{\text{intra}} \bar{X}_1 = 8,6 - (0,8421)(6) = 3,547$$

$$\alpha_2 = \bar{Y}_2 - \beta_{\text{intra}} \bar{X}_2 = 9,6 - (0,8421)(7,4) = 3,368$$

Una vez hallados estos parámetros, estimamos la *puntuación predicha*, Y'_{i1} , y el *error de predicción*, $(\epsilon_1)_{i1}$, correspondientes a cada sujeto, así como la *proporción de variabilidad no explicada por la covariable*, $(\epsilon_1)_{i1}^2$, en el primer grupo de tratamiento. Tales estimaciones se presentan en la Tabla 8.37.

TABLA 8.37 Cálculos basados en la regresión lineal para el desarrollo del ANCOVA (5): ejemplo práctico ($j = 1$)

X	Y	$Y'_{i1} = \alpha_1 + \beta_{\text{intra}} X_{i1}$	$(\epsilon_1)_{i1} = Y_{i1} - Y'_{i1}$	$(\epsilon_1)_{i1}^2$
7	9	9,44	- 0,44	0,194
6	10	8,60	1,40	1,960
4	7	6,92	0,08	0,006
8	9	10,28	- 1,28	1,638
5	8	7,76	0,24	0,058
30	43	SUMATORIOS (Σ)	0	3,856

Por último, elaboramos la Tabla 8.38, correspondiente a las estimaciones referidas al segundo grupo de tratamiento.

TABLA 8.38 Cálculos basados en la *regresión lineal* para el desarrollo del ANCOVA (6): ejemplo práctico ($j = 2$)

X	Y	$Y'_{i2} = \alpha_2 + \beta_{\text{intra}} X_{i2}$	$(\epsilon_1)_{i2} = Y_{i2} - Y'_{i2}$	$(\epsilon_1)_{i2}^2$
8	12	10,1	1,9	3,61
7	8	9,26	-1,26	1,59
6	7	8,42	-1,42	2,02
9	11	10,95	0,05	0,002
7	10	9,26	0,74	0,55
37	48	SUMATORIOS (Σ)	-0,01	7,772

TERCERA ETAPA

Llegados a este punto, podemos estimar la *razón F* para llevar a cabo la prueba de la hipótesis. En primer lugar, presentamos la fórmula para el cálculo de dicho estadístico y, posteriormente, la desarrollamos utilizando como referente los datos de nuestro ejemplo práctico.

$$F = \frac{\left[\sum_i^n (\epsilon_0)_{ij}^2 - \sum_i^n (\epsilon_1)_{ij}^2 \right]}{\frac{\sum_i^n (\epsilon_1)_{ij}^2}{a(n-1) - C}} \quad (8.108)$$

El error predicho bajo el modelo de la hipótesis nula (SCT_Y^*) (véase la Tabla 8.34) es:

$$\sum (\epsilon_0)_{ij}^2 = 11,6813$$

Por otra parte, el error predicho bajo el modelo de la hipótesis alternativa (SCE_Y^*) adopta el siguiente valor:

$$\sum (\epsilon_1)_{ij}^2 = \sum_j \sum_i (Y_{ij} - Y'_{ij})^2 = (\epsilon_1)_{i1}^2 + (\epsilon_1)_{i2}^2 = 3,856 + 7,772 = 11,628$$

Restando el segundo de tales errores al primero, calculamos el *nivel de mejora obtenido en la predicción* (SC_{entre}^*).

$$\sum (\epsilon_0)_{ij}^2 - \sum (\epsilon_1)_{ij}^2 = 11,6813 - 11,628 = 0,053$$

Llegados a este punto, podemos calcular el estadístico F aplicando la Fórmula (8.108):

$$F = \frac{\frac{[11,6813 - 11,628]}{(2 - 1)}}{\frac{11,628}{(2(5 - 1) - 1)}} = \frac{0,053}{1,661} = 0,03$$

Como cabe observar, el valor de la razón F obtenido mediante este segundo procedimiento es igual al estimado aplicando el primer procedimiento. Dado que la interpretación de los resultados del ANCOVA ya ha sido realizada previamente, no reincidiremos en tal aspecto.

Aunque en el presente ejemplo práctico carezca de utilidad, es importante señalar que, cuando se rechaza la hipótesis nula del ANCOVA, resulta muy conveniente calcular, en cada uno de los grupos de tratamiento, la *puntuación media ajustada al efecto de la covariable* ($\bar{Y}_j^*$), ya que dicho ajuste permite eliminar el sesgo sistemático derivado de la influencia que ejerce la covariable sobre la variable dependiente. Calcularemos, con un fin meramente didáctico, tales medias. Para ello, utilizaremos la Fórmula (8.109):

$$\bar{Y}_j^* = \bar{Y}_j - \beta_{\text{intra}}(\bar{X}_j - \bar{X}) \quad (8.109)$$

donde:

$\bar{Y}_j$ = Promedio de las puntuaciones obtenidas por los sujetos en la variable dependiente bajo el j -ésimo nivel de la variable de tratamiento.

β_{intra} = Pendiente media de regresión intragrupo.

$\bar{X}_j$ = Promedio de las puntuaciones obtenidas por los sujetos en la covariable bajo el j -ésimo nivel de la variable de tratamiento.

$\bar{X}$ = Promedio de las puntuaciones obtenidas por los sujetos en la covariable.

Aplicando la Fórmula (8.109) a los datos de nuestro ejemplo práctico:

$$\bar{Y}_j^* = \begin{pmatrix} 8,6 \\ 9,6 \end{pmatrix} - 0,8421 \left[\begin{pmatrix} 6 \\ 7,4 \end{pmatrix} - \begin{pmatrix} 6,7 \\ 6,7 \end{pmatrix} \right] = \begin{pmatrix} 9,18 \\ 9,01 \end{pmatrix}$$

Por último, solo resta decir que, en el caso de rechazar la hipótesis de nulidad del ANCOVA, en un diseño que consta de más de dos condiciones de tratamiento, deben aplicarse *contrastes a posteriori entre pares de medias ajustadas* a fin de determinar entre qué grupos existen diferencias estadísticamente significativas. Aunque en el presente texto no vamos a profundizar en el procedimiento, cabe señalar que, dado que las covariables se consideran, habitualmente, variables de efectos aleatorios, el valor del estadístico obtenido en la prueba de comparaciones múltiples se compara con la distribución de rango estudentizado generalizado de Bryan y Paulson (1976). El lector interesado en el desarrollo de dicho procedimiento, puede consultar, entre otros, el trabajo de Bryan y Paulson (1976), así como los textos de Maxwell y Delaney (1990) y Pascual, García y Frías (1995).

8.4.3.4. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

COMPROBACIÓN DE LOS SUPUESTOS DEL ANCOVA

Supuesto de independencia entre la variable independiente y la covariable

Para llevar a cabo la comprobación de este supuesto, debemos realizar un análisis de la varianza tomando, como variable independiente, la hora en la que se imparte la docencia y, como variable dependiente, las puntuaciones obtenidas por los sujetos en la covariable (pretest). Para ello, elegimos la opción **Anova de un factor** y realizamos el análisis tal y como se ha explicado en el Epígrafe 6.1.2.2 del Capítulo 6. La sintaxis para este análisis es la siguiente:

```
ONEWAY
  pretest BY hora
  /MISSING ANALYSIS.
```

Los resultados obtenidos mediante este análisis pueden observarse en la siguiente tabla:

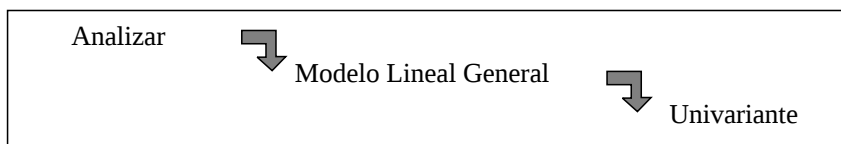
ANOVA de un factor

PRETEST Pretest		ANOVA			
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	4,900	1	4,900	2,579	0,147
Intra-grupos	15,200	8	1,900		
Total	20,100	9			

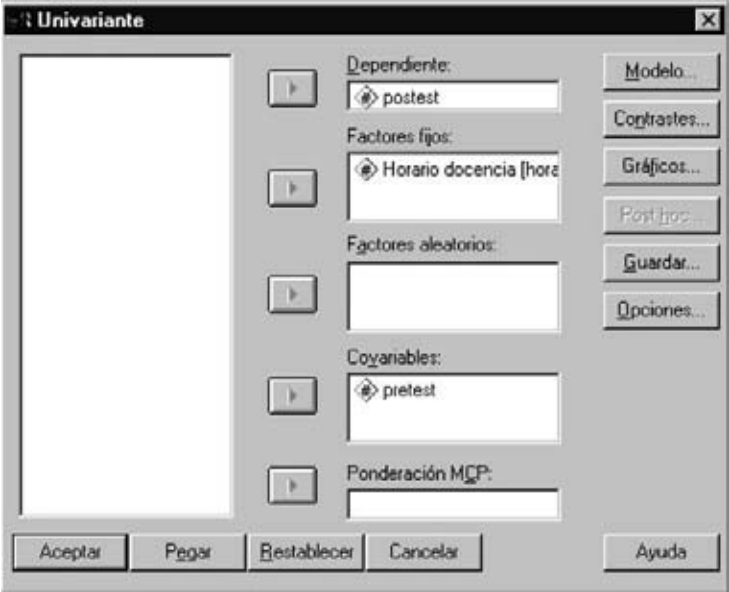
Supuesto de homogeneidad de las pendientes de regresión

La comprobación de este supuesto requiere verificar si la interacción entre la covariable y el tratamiento es, o no, estadísticamente significativa. Para ello:

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**, tal y como se ha expuesto en el Epígrafe 7.3.2.4 del Capítulo 7.



- En el siguiente cuadro de diálogo, indicamos la variable dependiente (en nuestro ejemplo la hemos denominado «postest»), la variable independiente (en nuestro ejemplo se trata de un factor fijo llamado «horario de docencia») y la covariable (en nuestro ejemplo la hemos denominado «pretest»).



- El análisis de la covarianza que realiza el SPSS 10.0 por defecto, no contempla el efecto de interacción entre el tratamiento y la covariable. Por ello, debemos *personalizar el modelo* escogiendo, en el cuadro de diálogo anterior, el menú **Modelo** y construir un modelo que contemple los efectos principales del tratamiento y de la covariable así como la interacción entre ambos factores. Este último término se obtiene seleccionando tales variables y pulsando, seguidamente, la opción *Interacción* del menú desplegable *Construir términos*. Tanto los efectos principales como el efecto de interacción deben colocarse en el cuadro derecho donde se indica *Modelo*, utilizando el botón destinado a tal fin. (El procedimiento para construir un modelo personalizado puede consultarse en el Epígrafe 8.2.2.4 del Capítulo 8).



- La sintaxis del análisis de la covarianza correspondiente a nuestro ejemplo, tras seguir los pasos arriba descritos, sería:

UNIANOVA

```
postest BY hora WITH pretest
/METHOD = SSTYPE(1)
/INTERCEPT = INCLUDE
/CRITERIA = ALPHA(.05)
/DESIGN = pretest hora hora*pretest.
```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
HORA Horario de docencia	1,00	Mañana	5
	2,00	Tarde	5

Pruebas de los efectos intersujetos

Variable dependiente: POSTEST Postest

Fuente	Suma de cuadrados tipo I	gl	Media cuadrática	F	Sig.
Modelo corregido	16,700 ^a	3	5,567	4,073	0,068
Intersección	828,100	1	828,100	605,927	0,000
PRETEST	13,218	1	13,218	9,672	0,021
HORA	6,054E-02 ³	1	6,054E-02	0,044	0,840
HORA * PRETEST	3,421	1	3,421	2,503	0,165
Error	8,200	6	1,367		
Total	853,000	10			
Total corregida	24,900	9			

^a R cuadrado = 0,671 (R cuadrado corregida = 0,506).

Como puede comprobarse en la tabla anterior, los valores de los estadísticos referidos a la interacción entre la variable independiente y la covariable son idénticos a los obtenidos mediante el cálculo manual presentado en la Tabla 8.28 de este capítulo.

$$^3 XE - 0n = x \cdot 10^{-n} = x \cdot \frac{1}{10^n}$$

En el caso que nos ocupa:

$$6,054E-02 = 6,054 \cdot (10^{-2}) = 6,054 \cdot \frac{1}{10^2} = 0,06054$$

ANÁLISIS DE LA COVARIANZA

- Retomamos el análisis previamente presentado (comprobación del supuesto de homogeneidad de las pendientes de regresión) y realizamos los dos primeros pasos. Una vez definidas la variable dependiente, la variable independiente y la covariable, el menú *Opciones* brinda la posibilidad de *mostrar las medias marginales* para el factor «horario de docencia», opción que proporciona las puntuaciones medias de los grupos ajustadas al efecto de la covariable. Asimismo, este menú permite seleccionar, de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo se han escogido las opciones *Mostrar* «estimaciones del tamaño del efecto» y «potencia observada»). Elegimos la opción *Aceptar*, obviando el tercer paso anteriormente expuesto referido a la especificación del modelo, es decir, el modelo en este caso es un modelo factorial completo.
- La sintaxis del análisis de la covarianza correspondiente a nuestro ejemplo, tras seguir los pasos arriba descritos, sería:

```
UNIANOVA
  postest BY hora WITH pretest
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /EMMEANS = TABLES(hora) WITH(pretest = MEAN)
  /PRINT = ETASQ OPOWER
  /CRITERIA = ALPHA(.05)
  /DESIGN = pretest hora.
```

- Resultados

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
HORA Horario de docencia	1,00	Mañana	5
	2,00	Tarde	5

Pruebas de los efectos intersujetos

Variable dependiente: POSTEST Postest

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	13,279 ^b	2	6,639	3,999	0,069	0,533	7,999	0,516
Intersección	3,916	1	3,916	2,359	0,168	0,252	2,359	0,265
PRETEST	10,779	1	10,779	6,493	0,038	0,481	6,493	0,592
HORA	6,054E-02	1	6,05E-02	0,036	0,854	0,005	0,036	0,053
Error	11,621	7	1,660					
Total	853,000	10						
Total corregida	24,900	9						

^a Calculado con alfa = 0,05.

^b R cuadrado = 0,533 (R cuadrado corregida = 0,400).

Medias marginales estimadas**Horario de docencia****Variable dependiente: POSTEST Postest**

Horario de docencia	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
1,00 Mañana	9,189 ^a	0,621	7,721	10,658
2,00 Tarde	9,011 ^a	0,621	7,542	10,479

^a Evaluado respecto a cómo aparecen las covariables en el modelo: PRETEST Pretest = 6,7000.

Como puede comprobarse en la tabla anterior, los valores de los estadísticos referidos al efecto de la variable independiente (una vez ajustado el efecto de la covariable) son idénticos a los obtenidos mediante el cálculo manual presentado en la Tabla 8.32 de este capítulo.

Asimismo, debido a que el componente de error de este análisis está ajustado al efecto de la covariable, la varianza debida al pretest es el componente residual de la variable dependiente que se halla determinado por la relación que ésta mantiene con la covariable (véase la descomposición de las sumas cuadráticas de la variable dependiente en la Tabla 8.24). Así, el estadístico *F* asociado a esta fuente de variación nos permite comprobar si *la relación existente entre la variable dependiente y la covariable es o no estadísticamente significativa* (véase la Tabla 8.25).

9

DISEÑOS EXPERIMENTALES DE MEDIDAS REPETIDAS

9.1. DISEÑOS SIMPLES Y FACTORIALES DE MEDIDAS TOTALMENTE REPETIDAS

9.1.1. Características generales del diseño de medidas repetidas

Los *diseños de medidas repetidas*, conocidos también como *diseños intrasujeto* y *diseños de medida múltiple*, se caracterizan por el registro de diversas medidas de la variable dependiente en un mismo grupo de sujetos. De este modo, las comparaciones entre las respuestas de los sujetos, ante los distintos tratamientos, se llevan a cabo dentro de un único grupo de sujetos (*comparaciones intrasujeto*), no estableciéndose comparaciones entre diferentes grupos de sujetos (*comparaciones intersujetos o intergrupos*). Como señala Arnau (1995e), en *contextos no experimentales*, característicos de la estrategia longitudinal, el interés por el diseño intrasujeto radica en la posibilidad que éste brinda de registrar un conjunto de puntuaciones o medidas de una variable, en dos o más puntos en el tiempo. Así, dado que desde una perspectiva longitudinal las respuestas de cada uno de los sujetos son función del tiempo, el diseño de medidas repetidas constituye un instrumento muy útil para la modelización de las curvas de crecimiento y la evaluación de los procesos de cambio en contextos evolutivos, sociales y educativos. No obstante, desde la perspectiva que aquí nos ocupa, a saber, desde el *enfoque experimental*, el principal objetivo del diseño intrasujeto consiste en estimar la efectividad de una serie sucesiva de tratamientos aplicados secuencialmente a las mismas unidades de observación.

Los diseños de medidas repetidas presentan varias **ventajas** con respecto a los diseños de medida única. A continuación destacamos las más relevantes:

- En primer lugar, constituyen instrumentos excelentes para reducir la varianza de error y minimizar la varianza sistemática secundaria. Dado que, en la estrategia intrasujeto,

el propio sujeto se convierte en criterio de bloqueo o de control, se extrae de la variabilidad del error una de sus principales fuentes, a saber, la varianza procedente de las diferencias individuales. De esta manera, el diseño intrasujeto consigue mayor precisión que cualquier otro tipo de diseño en la estimación de los efectos experimentales. De acuerdo con esta cualidad, diversos autores afirman que los diseños de medidas repetidas poseen mayor potencia estadística que los diseños completamente aleatorios (Arnau, 1995e; Riba, 1990; Stevens, 1992).

- Una segunda ventaja que, a nivel práctico, tiene gran importancia, radica en la menor cantidad de sujetos que se requieren para llevar a cabo el estudio con respecto a las investigaciones en las que se utilizan diseños de grupos totalmente al azar (Arnau, 1986; Pascual, 1995c; Vallejo, 1991). Maxwell y Delaney (1990) proporcionan una serie de tablas en las que se comparan los sujetos necesarios en la estrategia intrasujeto y en la estrategia intersujetos en función de la potencia estadística, del tamaño del efecto y de la cantidad de tratamientos que se les administran a los sujetos.

No obstante, el uso del diseño de medidas repetidas no está exento de ciertos **inconvenientes** relacionados con las características intrínsecas del propio diseño. Entre tales inconvenientes destacamos los siguientes:

- El primer problema asociado a los diseños intrasujeto se deriva de la secuencialidad con la que se aplican los tratamientos y da lugar a dos tipos principales de fuentes de confusión: los *efectos de período* y los *efectos residuales* o *efectos carry-over*. Los *efectos de período* hacen referencia al sesgo derivado del orden en el que se les administran los tratamientos a los sujetos. Normalmente, estos efectos se deben a factores, tales como el aprendizaje y la familiaridad (*efectos de período positivos*) o como la frustración y la fatiga (*efectos de período negativos*), que se desarrollan durante el intervalo de tiempo comprendido entre la administración del primer tratamiento y del último. Son factores directamente relacionados con el paso del tiempo y que ejercen una influencia selectiva sobre los tratamientos presentados en diferentes momentos, dentro del experimento. Por otra parte, los *efectos residuales* o *efectos carry-over* surgen cuando el efecto de un tratamiento no ha desaparecido en el momento en el que se introduce el siguiente tratamiento, es decir, cuando dicho efecto persiste una vez acabado el período de tratamiento.

Cabe señalar que existen diseños experimentales especialmente adecuados para neutralizar y estimar esta clase de efectos. Entre tales diseños cabe destacar los *diseños cross-over*, *alternativos* o *conmutativos* y los *diseños de cuadrado latino intrasujeto*. Este tipo de diseños controla los efectos de período y los efectos residuales, utilizando la estrategia consistente en *contrabalancear* las diferentes secuencias de tratamientos a través de los sujetos o a través de los grupos. Además de controlar tales fuentes de sesgo, estas estructuras de investigación también permiten estimar, de manera precisa, el efecto debido al orden o a la secuenciación de los tratamientos.

- Otro de los inconvenientes de los diseños intrasujeto, que afecta seriamente a la validez de conclusión estadística, hace referencia a la posible *correlación o dependencia entre las distintas puntuaciones de los sujetos*. Dado que son los mismos sujetos los que reciben cada una de las condiciones experimentales, cabe esperar que sus puntuaciones estén correlacionadas y que los errores no sean independientes entre sí, con lo que se incumple el supuesto más importante del ANOVA, a saber, el *supuesto de independencia entre las observaciones*. Teniendo en cuenta que el ANOVA no es robusto a la violación de este supuesto, la estimación de los efectos mediante el modelo de aná-

lisis de la varianza mixto está sujeta a error. Este tema será tratado posteriormente al abordar el *supuesto de esfericidad o circularidad* que debe cumplir todo diseño de medidas repetidas para el uso adecuado del estadístico F (véase el Epígrafe 9.1.2, dentro de este mismo epígrafe).

- Como tercer inconveniente, Pascual (1995c) plantea la existencia de, al menos, una variable de naturaleza aleatoria (la *variable sujeto*) dentro del diseño. La presencia de este factor implica que los diseños de medidas repetidas son necesariamente *modelos de efectos aleatorios* o *modelos mixtos*. Debido a la naturaleza cambiante de la variable aleatoria, es presumible la existencia de un efecto de interacción entre dicha variable y la variable fija, con lo que la variabilidad entre las medias del efecto fijo tiende a incrementarse y, por tanto, aumenta la probabilidad de cometer un error de tipo I en la estimación de los parámetros.

9.1.2. El análisis de datos en los diseños de medidas totalmente repetidas

9.1.2.1. Supuestos básicos para el análisis y alternativas ante su incumplimiento

El modelo analítico utilizado habitualmente para llevar a cabo la **prueba de la hipótesis** con este tipo de diseños se conoce como *análisis de la varianza mixto* (Bock, 1975; Kirk, 1982; Winer, 1971). Si consideramos el *formato más simple del diseño de medidas repetidas*, el modelo matemático no aditivo que subyace al análisis de la varianza, bajo el supuesto de la hipótesis alternativa, responde a la siguiente expresión:

$$y_{ij} = \mu + \eta_i + \alpha_j + (\eta\alpha)_{ij} + \varepsilon_{ij} \quad (9.1)$$

donde:

- y_{ij} = Puntuación obtenida en la variable dependiente por el sujeto i bajo el j -ésimo nivel de la variable de tratamiento.
- μ = Media común a todas las observaciones.
- η_i = Componente específico asociado al sujeto i y constante a lo largo de las observaciones.
- α_j = Efecto debido a la administración del j -ésimo nivel de la variable de tratamiento.
- $(\eta\alpha)_{ij}$ = Efecto debido a la interacción entre el i -ésimo sujeto y el j -ésimo nivel de la variable de tratamiento.
- ε_{ij} = Componente de error específico asociado al sujeto i y al j -ésimo nivel de la variable de tratamiento.

Se asume que ε_{ij} es independiente de η_i y que los sujetos han sido seleccionados de una población donde el componente η_i (factor aleatorio) tiene una distribución independiente, definida por:

$$\eta_i \simeq NID(O, \sigma_\eta^2)$$

Se asume, también, que el componente de error tiene una distribución:

$$\varepsilon_{ij} \simeq NID(O, \sigma_\varepsilon^2)$$

y que los niveles del factor de tratamiento son fijos.

La presencia de un *factor aleatorio* en el modelo requiere considerar la interacción como determinante del efecto de la variable fija. En consecuencia, el modelo estadístico para la estimación de los parámetros debe asumir el término de interacción como un componente adicional de las medias cuadráticas esperadas del factor fijo. Por esta razón, en los diseños en los que se incluye alguna variable aleatoria, la interacción entre dicha variable y la variable de tratamiento se toma como componente de error (media cuadrática del error) para estimar el efecto de la variable de tratamiento. Dado que, como señala Pascual (1995c), en los modelos con componentes aleatorios los parámetros de interés incluyen la varianza asociada con las distribuciones de tales componentes, es muy importante identificarlos a fin de que la razón F no resulte sesgada debido a una selección inadecuada del término de error.

El adecuado ajuste del modelo de análisis de la varianza mixto requiere el cumplimiento del *supuesto de esfericidad* o *circularidad*, es decir, requiere que las varianzas de las diferencias entre cada par de medias de medidas repetidas sean constantes: $\sigma_{ja-jb}^2 = \sigma_{jc-jd}^2$, $\forall (a, b) \neq (c, d)$ (Harris, 1975; Huynh y Feldt, 1970; Kirk, 1982; McCall y Appelbaum, 1973; Mendoza, 1980; Rouanet y Lepine, 1970; Vallejo, 1991). Siempre que se cumple este supuesto, el cociente de medias cuadráticas sigue exactamente la distribución F , a pesar de que exista covarianza entre las observaciones. Una forma particular de esfericidad es la denominada *simetría compuesta* o *combinada*, cuya presencia exige que las correlaciones entre todos los pares de medidas repetidas sean iguales. Requiere la igualdad de las varianzas muestrales y la igualdad de las correlaciones (Maxwell y Delaney, 1990; Pascual y Camarasa, 1991). Aunque el incumplimiento de la simetría compuesta no implica necesariamente que se ha transgredido la esfericidad, indica que probablemente se ha producido tal transgresión. Como ya demostró Box (1954) en la década de los cincuenta, en los casos en los que no se cumple el supuesto de esfericidad, se obtienen estimaciones positivamente sesgadas de la razón F .

Como cabe deducir de lo que acabamos de exponer, antes de aplicar el análisis de la varianza mixto, conviene verificar si se cumple el *supuesto de esfericidad de la matriz de varianzas-covarianzas del diseño*. Para ello, se puede utilizar la prueba de esfericidad de Mauchly (1940), aunque hemos de señalar que es una prueba sensible al incumplimiento del supuesto de normalidad (Keselman, Rogan, Mendoza y Breen, 1980; Maxwell y Delaney, 1990; Stevens, 1992). Por otra parte, el *supuesto de simetría combinada* puede verificarse aplicando la prueba de Box (1950). Arnau (1995e) describe de forma exhaustiva los cálculos que se deben realizar para aplicar ambas pruebas. Dado que en el presente texto no vamos a proceder a su desarrollo, remitimos al lector interesado, a la obra de Arnau.

En caso de que no se cumpla el supuesto de esfericidad, se pueden adoptar tres *alternativas*:

- Simplificar los grados de libertad asociados al numerador $(k - 1)$ y al denominador $(n - 1)(k - 1)$ de la razón entre varianzas, de forma que la F sea mucho más conservadora.
- Corregir los grados de libertad de la razón F mediante la $\hat{\epsilon}$ de Greenhouse y Geisser (1959).
- Aplicar el análisis multivariante de la varianza (MANOVA).

Las dos primeras alternativas son estrategias univariadas, mediante las que se llevan a cabo determinadas correcciones en el estadístico F , con el objetivo de reducir la probabilidad de cometer un error de tipo I en la decisión estadística. Ambas se engloban bajo la denominación de *modelo mixto univariante*. La tercera alternativa requiere el cumplimiento de menos supuestos en el análisis y se conoce como *modelo multivariante*. Maxwell y Delaney

(1990) afirman que el modelo mixto es más potente que el modelo multivariante, siempre que se satisfaga el supuesto de esfericidad. Cuando este supuesto no se cumple, el modelo mixto tiene mayor potencia que el multivariante con muestras pequeñas de sujetos, mientras que el modelo multivariante es más potente, a medida que aumenta el número de unidades experimentales (Riba, 1990). Davidson (1972) establece el criterio de elección entre ambas estrategias en torno a un tamaño muestral de $20 + k$ sujetos, siendo k el número de tratamientos. Maxwell y Delaney (1990) adoptan un criterio de decisión más flexible, cercano a un tamaño muestral de $10 + k$ sujetos.

A modo ilustrativo, supongamos que trabajamos con un diseño unifactorial intrasujeto, en el que la variable independiente consta de tres niveles ($k = 3$) que son administrados, secuencialmente, a 5 sujetos ($n = 5$) y en el que se incumple el supuesto de esfericidad. En tales circunstancias, la utilización de la primera de las alternativas citadas, a saber, la aplicación de la *prueba F conservadora*, consistiría en ajustar los grados de libertad del diseño original multiplicándolos por un factor corrector, ε , cuyo límite inferior o valor que permite corregir la máxima transgresión posible del supuesto de esfericidad es $\varepsilon = \frac{1}{k-1}$. De este

modo, en el diseño que nos ocupa, los grados de libertad corregidos mediante la prueba *F conservadora* adoptarían los siguientes valores:

Grados de libertad del numerador:

$$\frac{1}{k-1} (k-1) = \frac{1}{3-1} (3-1) = 1$$

Grados de libertad del denominador:

$$\frac{1}{k-1} (n-1)(k-1) = \frac{1}{3-1} (5-1)(3-1) = 4$$

Tras la corrección de los grados de libertad, el valor crítico para la prueba de la hipótesis sería $F_{0,95; 1,4} = 7,71$ que, como se puede comprobar, es superior al que se obtendría en caso de trabajar con los grados de libertad sin ajustar ($F_{0,95; 2,8} = 4,46$).

La prueba *F conservadora* es una prueba muy simple, dado que su aplicación no requiere sino utilizar, sistemáticamente, 1 y $n-1$ grados de libertad para las fuentes de varianza explicada y no explicada, respectivamente. Sin embargo, esta prueba es altamente conservadora y, en consecuencia, reduce la potencia de la prueba estadística.

Por ello, en la actualidad, la mayoría de los investigadores tienden a afrontar el problema del incumplimiento del supuesto de esfericidad adoptando la segunda de las alternativas citadas, a saber, corrigiendo los grados de libertad de la razón *F* mediante la $\hat{\varepsilon}$ de *Greenhouse* y *Geisser*. La lógica subyacente a la aplicación de esta prueba es similar a la de la prueba *F conservadora*. De hecho, en ambas pruebas, el objetivo del investigador consiste en ajustar los grados de libertad del diseño original de tal forma que, tras el ajuste, el valor crítico de la razón *F* sea mayor que el que se hubiera obtenido en caso de no haber llevado a cabo tal corrección. El ajuste de los grados de libertad de la razón *F* mediante la $\hat{\varepsilon}$ de *Greenhouse* y *Geisser* se realiza multiplicando los grados de libertad del diseño original por el siguiente factor corrector:

$$\hat{\varepsilon} \text{ de Greenhouse y Geisser} = \frac{k^2(\bar{S}_{jj} - \bar{S}_{..})^2}{(k-1)[\sum S_{jj}^2 - (2k \sum \bar{S}_j^2) + (k^2 \bar{S}_{..}^2)]} \quad (9.2)$$

donde:

- k = Cantidad de niveles de la variable de tratamiento.
- $\bar{S}_{jj}$ = Promedio de los elementos de la diagonal principal (es decir, promedio de las varianzas) de la matriz de varianzas-covarianzas poblacional.
- $\bar{S}_{..}$ = Promedio de todos los elementos de la matriz de varianzas-covarianzas poblacional.
- $\sum \sum S_{jj}^2$ = Suma al cuadrado de cada elemento de la matriz de varianzas-covarianzas poblacional.
- $\sum \bar{S}_j^2$ = Suma al cuadrado del promedio de los elementos de cada fila de la matriz de varianzas-covarianzas poblacional.

La corrección de los grados de libertad de la razón F mediante la $\hat{\varepsilon}$ de Greenhouse y Geisser permite controlar adecuadamente el error de tipo I fijado, a priori, por el investigador.

Por último, cabe señalar que la tercera de las alternativas citadas, es decir, la aplicación del *modelo multivariante* no requiere el cumplimiento del supuesto de esfericidad (Girden, 1992). Desde este enfoque, cada una de las medidas que se registran en cada sujeto de forma repetida se considera como una variable dependiente diferente. En la aproximación multivariada, se trabaja con los determinantes de las matrices de sumas de cuadrados, con lo que se solucionan los problemas asociados a las varianzas y covarianzas poblacionales. Además, se tiene en cuenta información relevante, como la referida a la covarianza existente entre las diferentes variables dependientes incluidas en el diseño.

Como señalan Pascual, Frías y García (1996), las diferentes alternativas univariadas o asociadas al modelo mixto univariante permiten ajustar los grados de libertad del diseño original, tratando de compensar el incremento en el error de tipo I que se produce bajo el incumplimiento del supuesto de esfericidad. Sin embargo, lo único que se asegura con tales procedimientos es que el α real tenga un valor *aproximado* al nivel de α nominal fijado, a priori, por el investigador. El modelo multivariante, por su parte, garantiza matemáticamente la *igualdad* entre el error de tipo I y el α nominal, siempre y cuando se cumpla el supuesto de normalidad multivariada. Los autores arriba citados desarrollan, de forma muy didáctica, una de las pruebas más utilizadas para verificar el cumplimiento del supuesto de esfericidad, a saber, la prueba de Mauchly, así como las principales estrategias que pueden adoptarse ante la violación de tal supuesto. El lector interesado en el desarrollo matemático de dichas pruebas puede consultarlo en la obra de Pascual, Frías y García (1996).

9.1.2.2. *Análisis de la varianza mixto para el diseño intrasujeto simple (diseño de tratamientos $\times$ sujetos)*

• **Modelo general de análisis**

El **diseño experimental intrasujeto simple**, conocido también como *diseño de tratamientos $\times$ sujetos*, es una estructura de investigación en la que se cruza una variable de tratamiento con una variable sujeto. Al igual que en el diseño factorial de dos factores intersujetos, el diseño consta de dos factores: el factor de tratamiento, que asume un modelo de efectos fijos, y el factor sujeto, asumiendo este último un modelo de efectos aleatorios.

Como ya se ha señalado en el subapartado anterior, la **prueba de la hipótesis**, en este tipo de diseño, se lleva a cabo aplicando el *análisis de la varianza mixto*. Sin embargo, la ecuación matemática subyacente a dicho análisis, bajo el supuesto de la hipótesis alternativa, puede ajustarse a dos tipos de modelos: el modelo aditivo y el modelo no aditivo.

El *modelo de aditividad de los efectos* responde a la siguiente expresión:

$$y_{ij} = \mu + \eta_i + \alpha_j + \varepsilon_{ij} \quad (9.3)$$

cuyos componentes ya han sido descritos en el Epígrafe 9.1.2.1. El supuesto específico del modelo aditivo es que, tanto η_i como α_j , contribuyen de forma aditiva a la puntuación del sujeto y que, por tanto, son independientes entre sí. No obstante, es muy frecuente que se produzca una interacción entre la variable de tratamiento y la variable sujeto, lo que significa que existe una porción de la puntuación observada que no es función aditiva, ni del efecto principal de sujeto, ni del efecto principal de tratamiento. En este caso, el modelo matemático que representa más adecuadamente los datos es el *modelo de no aditividad de los efectos*, que responde a la siguiente expresión:

$$y_{ij} = \mu + \eta_i + \alpha_j + (\eta\alpha)_{ij} + \varepsilon_{ij} \quad (9.1)$$

Los componentes de este modelo también han sido descritos en el Epígrafe 9.1.2.1.

Dado que cada combinación factorial consta de una sola observación, el efecto debido a la interacción $(\eta\alpha)_{ij}$ no se puede calcular independientemente del efecto del error experimental ε_{ij} . En consecuencia, el componente residual incluye tanto la variación debida al error, como la forma específica en la que reacciona cada sujeto ante los diferentes tratamientos. La prueba de Tukey (1949) permite detectar si existe interacción $(\eta\alpha)_{ij}$, es decir, permite saber si las desviaciones de las puntuaciones de los sujetos, con respecto a los promedios de los tratamientos, son diferentes de un sujeto a otro (el lector interesado en el desarrollo matemático de dicha prueba puede consultarla en el Epígrafe 8.1.2 referido a los *diseños de bloques aleatorios*).

Como señala Arnau (1986), la distinción entre los modelos aditivo y no aditivo es importante en relación con el supuesto básico de *homogeneidad de covarianzas* ($\text{cov}(y_{ij}, y'_{ij}) = \text{constante}$ para todos los pares j, j' , donde $j \neq j'$). Cuando la interacción del modelo es nula, la covarianza entre cualquier par de tratamientos es la misma y, por tanto, en el modelo aditivo siempre se cumple la condición de homogeneidad. Sin embargo, si los datos se ajustan al modelo de no aditividad, se puede transgredir tal supuesto. Como se ha señalado en el subapartado anterior, tanto la heterogeneidad de las varianzas como la de las covarianzas de un diseño cuyos datos están correlacionados, produce un sesgo positivo en la estimación de la F ordinaria, por lo que deben adoptarse estrategias de análisis alternativas para llevar a cabo la prueba de la hipótesis (véase el Epígrafe 9.1.2.1).

• Ejemplo práctico

Supongamos que, en el ámbito de los procesos psicológicos básicos, realizamos una investigación con el objetivo de examinar la memoria a corto plazo. En concreto, se pretende analizar la influencia que ejerce el *tiempo transcurrido entre la presentación visual de una lista de palabras y la realización de una prueba de recuerdo* (factor A) sobre la *cantidad de palabras de dicha lista que recuerda el sujeto*. A tal fin, se seleccionan tres intervalos temporales distintos: (a_1) *intervalo de 10 segundos*, (a_2) *intervalo de 20 segundos* y (a_3) *intervalo de 25 segundos*. Se escoge, al azar, una muestra de seis sujetos y se les presentan, sucesivamente y de forma aleatoria, diferentes listas de 15 palabras cada una. Tras la presentación se registra la cantidad de palabras recordadas por cada sujeto 10 segundos, 20 segundos y 25 segundos después de visualizar la lista. En la Tabla 9.1 pueden observarse los resultados obtenidos en el estudio.

TABLA 9.1
Matriz de datos del experimento

		A (Intervalo establecido para la prueba de recuerdo)			Cuadrados de las sumas	Medias marginales
		a_1 (10 sg)	a_2 (20 sg)	a_3 (25 sg)		
S (Sujetos)	S_1	7	4	5	256	5,33
	S_2	8	5	7	400	6,66
	S_3	8	5	3	256	5,33
	S_4	9	4	2	225	5
	S_5	10	3	6	361	6,33
	S_6	7	6	5	324	6
Cuadrados de las sumas		2.401	729	784	10.816	
Medias marginales		8,16	4,5	4,66		5,77

La Tabla 9.1 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan las *seis categorías del factor sujeto* (S_1, S_2, S_3, S_4, S_5 y S_6) y, en las columnas, los tres niveles del *factor de tratamiento* (a_1, a_2 y a_3). Por tanto, $i = 1, 2, 3, 4, 5, 6$ y $j = 1, 2, 3$.

Antes de abordar el análisis de la varianza examinaremos con más detalle la estructura de la tabla (véase la Tabla 9.2).

TABLA 9.2
Datos correspondientes a un diseño intrasujeto simple: modelo general

Variable independiente	a_1	...	a_j	...	a_a	$T^2_{i.}$	Medias marginales
s_1	Y_{11}	...	Y_{1j}	...	Y_{1a}	$T^2_{1.}$	$\bar{Y}_{1.}$
...		...		...			
s_i	Y_{i1}	...	Y_{ij}	...	Y_{ia}	$T^2_{i.}$	$\bar{Y}_{i.}$
...		...		...			
s_n	Y_{n1}	...	Y_{nj}	...	Y_{na}	$T^2_{n.}$	$\bar{Y}_{n.}$
$T^2_{.j}$	$T^2_{.1}$	...	$T^2_{.j}$	...	$T^2_{.a}$	$T^2_{..}$	
Medias marginales	$\bar{Y}_{.1}$		$\bar{Y}_{.j}$		$\bar{Y}_{.a}$		$\bar{Y}_{..}$

Cada una de las celdillas de la Tabla 9.2 corresponde a la *puntuación obtenida por un determinado sujeto bajo un determinado tratamiento* que se le ha administrado en un

momento concreto. En la parte derecha (en la columna T_i^2) se presentan los *cuadrados de las sumas correspondientes a cada sujeto* ($i = 1, 2, \dots, n$) y en la parte inferior de la tabla (en la fila T_j^2) los *cuadrados de las sumas correspondientes a cada tratamiento* ($j = 1, 2, \dots, a$). Como cabe apreciar, los primeros corresponden a los *niveles del factor S* y los segundos a las *categorías del factor A*. Con respecto a la notación, la letra T hace referencia a la suma de una serie de puntuaciones y, más específicamente, a la suma de los datos que se expresan mediante los subíndices asociados a dicha letra. Por ejemplo, el símbolo $T_{..}^2$ expresa el *cuadrado de la suma de todas las puntuaciones*. Por último, en los *márgenes* se presentan las *medias aritméticas* correspondientes a cada fila o columna.

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño y, suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 9.1. Como ya se habrá percatado el lector, la estructura del diseño que nos ocupa es similar a la del diseño factorial $A \times B$ al azar (véase el Capítulo 7).

• Desarrollo del análisis de la varianza mixto para el diseño intrasujeto simple

Al igual que en los diseños precedentes, calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la ecuación estructural del ANOVA. Es necesario señalar que nuestras estimaciones partirán del supuesto de no aditividad de los efectos, en cuyo caso la interacción $(\eta\alpha)_{ij}$ se considera como componente residual del modelo. En concreto, desarrollaremos el ANOVA tomando como referencia la siguiente expresión:

$$Y_{ij} = \mu + \eta_i + \alpha_j + (\eta\alpha)_{ij} \quad (9.4)$$

Procedimiento 1

En la Tabla 9.3 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas.

Tras obtener el valor de C , llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{3 \cdot 6} (7 + 8 + 8 + 9 + 10 + 7 + 4 + 5 + 5 + 4 + 3 + 6 + 5 + 7 + 3 + 2 + 6 + 5)^2 = \frac{(104)^2}{18} = 600,88$$

Una vez obtenido este valor, procedemos a calcular la SCA , la SCS , la SCT y la $SCAS$, respectivamente:

$$SCA = \frac{1}{6} [(49)^2 + (27)^2 + (28)^2] - 600,88 = 51,45$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{.j}$, a saber, a los elementos de la fila T_j^2 de la Tabla 9.2.

$$SCS = \frac{1}{3} [(16)^2 + (20)^2 + (16)^2 + (15)^2 + (19)^2 + (18)^2] - 600,88 = 6,45$$

TABLA 9.3 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño intrasujeto simple bajo el supuesto de no aditividad de los efectos

Factor de tratamiento (A)	$SCA = \left(\frac{1}{n} \sum_j T_{.j}^2 \right) - C = \left[\frac{1}{n} \sum_j \left(\sum_i Y_{ij} \right)^2 \right] - C$ (9.5)
Factor sujeto (S)	$SCS = \left(\frac{1}{a} \sum_i T_{i.}^2 \right) - C = \left[\frac{1}{a} \sum_i \left(\sum_j Y_{ij} \right)^2 \right] - C$ (9.6)
Interacción A × S (error)	$SCAS = \left(\left[\sum_j \sum_i Y_{ij}^2 \right] - C \right) - (SCA + SCS)$ (9.7)
Variabilidad total	$SCT = \sum_j \sum_i Y_{ij}^2 - C$ (9.8)
C	$C = \frac{1}{N} T_{..}^2 = \frac{1}{an} \left(\sum_j \sum_i Y_{ij} \right)^2$ (9.9)

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{i.}$, es decir, a los elementos de la columna $T_{i.}^2$ de la Tabla 9.2.

$$SCT = \left[(7)^2 + (8)^2 + (8)^2 + (9)^2 + (10)^2 + (7)^2 + (4)^2 + (5)^2 + (5)^2 + \right. \\ \left. + (4)^2 + (3)^2 + (6)^2 + (5)^2 + (7)^2 + (3)^2 + (2)^2 + (6)^2 + (5)^2 \right] - 600,88 = 81,12$$

Una vez calculada la variabilidad total, procedemos a estimar la suma cuadrática correspondiente a la interacción entre el factor A y el factor S.

$$SCAS = SCT - (SCA + SCS) = 81,12 - (51,45 + 6,45) = 23,22$$

Procedimiento 2: Desarrollo mediante vectores

En primer lugar, hallaremos los vectores asociados a cada uno de los parámetros de la ecuación estructural del ANOVA correspondiente al diseño intrasujeto simple (Fórmula 9.4) y, posteriormente, calcularemos las sumas de cuadrados de los diferentes efectos.

Omitiendo el cálculo del *vector Y* que, como es sabido, es el *vector columna* de todas las *puntuaciones directas*, comenzaremos calculando los valores correspondientes a los vectores asociados a los efectos principales de los factores A y S.

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada una de las categorías del factor A para estimar, a partir de tales medias, los parámetros asociados al efecto del factor de tratamiento A, α_j .

$$\mu = \frac{1}{a \cdot n} \sum_j \sum_i Y_{ij} \quad (9.10)$$

VECTOR S

A fin de obtener la variabilidad correspondiente al factor sujeto, estimaremos los parámetros asociados a dicho factor, π_i . Para ello, comenzaremos calculando los promedios de las categorías del factor S.

$$\pi_i = \mu_{i.} - \mu \quad (9.13)$$

$$\mu_{i.} = \frac{1}{a} \sum_j Y_{ij}$$

respectivamente:

$$\mu_{1.} = \frac{1}{3} (7 + 4 + 5) = 5,33$$

$$\mu_{2.} = \frac{1}{3} (8 + 5 + 7) = 6,66$$

$$\mu_{3.} = \frac{1}{3} (8 + 5 + 3) = 5,33$$

$$\mu_{4.} = \frac{1}{3} (9 + 4 + 2) = 5,00$$

$$\mu_{5.} = \frac{1}{3} (10 + 3 + 6) = 6,33$$

$$\mu_{6.} = \frac{1}{3} (7 + 6 + 5) = 6,00$$

Por tanto:

$$\pi_1 = \mu_{1.} - \mu = 5,33 - 5,77 = -0,44$$

$$\pi_2 = \mu_{2.} - \mu = 6,66 - 5,77 = 0,89$$

$$\pi_3 = \mu_{3.} - \mu = 5,33 - 5,77 = -0,44$$

$$\pi_4 = \mu_{4.} - \mu = 5,00 - 5,77 = -0,77$$

$$\pi_5 = \mu_{5.} - \mu = 6,33 - 5,77 = 0,56$$

$$\pi_6 = \mu_{6.} - \mu = 6,00 - 5,77 = 0,23$$

El *vector* S adopta los siguientes valores:

$$S = \{\pi\} = \begin{bmatrix} -0,44 \\ 0,89 \\ -0,44 \\ -0,77 \\ 0,56 \\ 0,23 \\ -0,44 \\ 0,89 \\ -0,44 \\ -0,77 \\ 0,56 \\ 0,23 \\ -0,44 \\ 0,89 \\ -0,44 \\ -0,77 \\ 0,56 \\ 0,23 \end{bmatrix} \quad (9.14)$$

VECTOR $\{\pi\alpha\}$ O RESIDUAL

La variabilidad asociada al vector $\{\pi\alpha\}$ es la correspondiente al efecto de interacción entre el factor sujeto y el factor tratamiento. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\pi\alpha)_{ij}$.

$$(\pi\alpha)_{ij} = Y_{ij} - (\mu + \pi_i + \alpha_j) \quad (9.15)$$

Aplicando la Fórmula (9.15) a los datos de nuestro ejemplo obtenemos:

$$\begin{aligned} (\pi\alpha)_{11} &= Y_{11} - (\mu + \pi_1 + \alpha_1) = 7 - (5,77 + (-0,44) + 2,39) = -0,72 \\ (\pi\alpha)_{12} &= Y_{12} - (\mu + \pi_1 + \alpha_2) = 4 - (5,77 + (-0,44) + (-1,27)) = -0,06 \\ (\pi\alpha)_{13} &= Y_{13} - (\mu + \pi_1 + \alpha_3) = 5 - (5,77 + (-0,44) + (-1,11)) = 0,78 \\ (\pi\alpha)_{21} &= Y_{21} - (\mu + \pi_2 + \alpha_1) = 8 - (5,77 + 0,89 + 2,39) = -1,05 \\ (\pi\alpha)_{22} &= Y_{22} - (\mu + \pi_2 + \alpha_2) = 5 - (5,77 + 0,89 + (-1,27)) = -0,39 \\ (\pi\alpha)_{23} &= Y_{23} - (\mu + \pi_2 + \alpha_3) = 7 - (5,77 + 0,89 + (-1,11)) = 1,45 \\ (\pi\alpha)_{31} &= Y_{31} - (\mu + \pi_3 + \alpha_1) = 8 - (5,77 + (-0,44) + 2,39) = 0,28 \\ (\pi\alpha)_{32} &= Y_{32} - (\mu + \pi_3 + \alpha_2) = 5 - (5,77 + (-0,44) + (-1,27)) = 0,94 \\ (\pi\alpha)_{33} &= Y_{33} - (\mu + \pi_3 + \alpha_3) = 3 - (5,77 + (-0,44) + (-1,11)) = -1,22 \\ (\pi\alpha)_{41} &= Y_{41} - (\mu + \pi_4 + \alpha_1) = 9 - (5,77 + (-0,77) + 2,39) = 1,61 \\ (\pi\alpha)_{42} &= Y_{42} - (\mu + \pi_4 + \alpha_2) = 4 - (5,77 + (-0,77) + (-1,27)) = 0,27 \\ (\pi\alpha)_{43} &= Y_{43} - (\mu + \pi_4 + \alpha_3) = 2 - (5,77 + (-0,77) + (-1,11)) = -1,89 \\ (\pi\alpha)_{51} &= Y_{51} - (\mu + \pi_5 + \alpha_1) = 10 - (5,77 + 0,56 + 2,39) = 1,28 \\ (\pi\alpha)_{52} &= Y_{52} - (\mu + \pi_5 + \alpha_2) = 3 - (5,77 + 0,56 + (-1,27)) = -2,06 \\ (\pi\alpha)_{53} &= Y_{53} - (\mu + \pi_5 + \alpha_3) = 6 - (5,77 + 0,56 + (-1,11)) = 0,78 \end{aligned}$$

$$(\pi\alpha)_{61} = Y_{61} - (\mu + \pi_6 + \alpha_1) = 7 - (5,77 + 0,23 + 2,39) = -1,39$$

$$(\pi\alpha)_{62} = Y_{62} - (\mu + \pi_6 + \alpha_2) = 6 - (5,77 + 0,23 + (-1,27)) = 1,27$$

$$(\pi\alpha)_{63} = Y_{63} - (\mu + \pi_6 + \alpha_3) = 5 - (5,77 + 0,23 + (-1,11)) = 0,11$$

El vector $\{\pi\alpha\}$ adopta los siguientes valores:

$$\{\pi\alpha\} = \begin{bmatrix} -0,72 \\ -1,05 \\ 0,28 \\ 1,61 \\ 1,28 \\ -1,39 \\ -0,06 \\ -0,39 \\ 0,94 \\ 0,27 \\ -2,06 \\ 1,27 \\ 0,78 \\ 1,45 \\ -1,22 \\ -1,89 \\ 0,78 \\ 0,11 \end{bmatrix} \quad (9.16)$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = n \sum_j \alpha_j^2 = 6[(2,39)^2 + (-1,27)^2 + (-1,11)^2] = 51,34 \quad (9.17)$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR S (SCS):

$$SCS = a \sum_i \pi_i^2 = 3[(-0,44)^2 + (0,89)^2 + (-0,44)^2 + (-0,77)^2 + (0,56)^2 + (0,23)^2] \quad (9.18)$$

$$SCS = 6,41$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y S O RESIDUAL (SCAS):

$$SCAS = \sum_j \sum_i (\pi\alpha)^2 \quad (9.19)$$

$$SCAS = \begin{bmatrix} (-0,72)^2 + (-0,06)^2 + (0,78)^2 + (-1,05)^2 + (-0,39)^2 + (1,45)^2 + \\ + (0,28)^2 + (0,94)^2 + (-1,22)^2 + (1,61)^2 + (0,27)^2 + (-1,89)^2 + \\ + (1,28)^2 + (-2,06)^2 + (0,78)^2 + (-1,39)^2 + (1,27)^2 + (0,11)^2 \end{bmatrix} = 23,22$$

VARIABILIDAD TOTAL (SCT):

$$SCT = 23,22 + 6,41 + 51,34 = 81,12 \quad (9.20)$$

Multiplicando cada vector por su traspuesto, obtenemos los mismos resultados:

$$SCA = \{\alpha\}^T \{\alpha\} = 51,45 \quad (9.21)$$

$$SCS = \{\pi\}^T \{\pi\} = 6,41 \quad (9.22)$$

$$SCAS = \{\pi\alpha\}^T \{\pi\alpha\} = 23,22 \quad (9.23)$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 9.4 Análisis de la varianza mixto para un diseño intrasujeto simple bajo el supuesto de no aditividad de los efectos: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	<i>F</i>
Factor de tratamiento (factor A)	SCA = 51,45	$a - 1 = 2$	MCA = 25,725	$F_A = 11,079$
Factor sujeto (factor S)	SCS = 6,41	$n - 1 = 5$	MCS = 1,282	$F_S = 0,552$
Interacción A × S (error)	SCAS = 23,22	$(a - 1)(n - 1) = 10$	MCAS = 2,322	
Total	SCT = 81,12	$an - 1 = 17$		

Tras la obtención de las *F* observadas, debemos determinar si la variabilidad explicada por los factores es o no significativa. Para ello, recurrimos a las *tablas de los valores críticos de la distribución F*. Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos.

TABLA 9.5 Comparación entre las *F* observadas y las *F* teóricas con un nivel de confianza del 95 %

Fuente de variación	<i>F</i> crítica _(0,95; gl1/gl2)	<i>F</i> observada	Diferencia
Factor A	$F_{0,95; 2/10} = 4,10$	$F_A = 11,079$	$4,10 < 11,079$
Factor S	$F_{0,95; 5/10} = 3,33$	$F_S = 0,552$	$3,33 > 0,552$

Como puede apreciarse en la Tabla 9.5, rechazamos la hipótesis nula para el efecto del *factor tratamiento* o *factor A*. Por tanto, cabe concluir que el *intervalo de tiempo transcurrido entre la presentación de la lista de palabras y la realización de la prueba de recuerdo* ejerce una influencia estadísticamente significativa ($p < 0,05$) sobre la *cantidad de palabras de dicha lista que recuerda el sujeto*. Por otra parte, el *factor sujeto* aporta al modelo una fuente de variación cuya influencia sobre la variable criterio no resulta estadísticamente significativa. Sin embargo, hemos de señalar que la prueba de la hipótesis nula para esta última fuente de variación suele carecer de interés. En este sentido, Riba (1990) afirma que, dado que el objetivo del investigador no radica en examinar los efectos individuales de cada uno de los sujetos, no es usual llevar a cabo una estimación de los parámetros asociados a dicho factor. De hecho, en el contexto de la psicología experimental básica y, en el de las ciencias del comportamiento en general, la variabilidad asociada a las diferencias individuales tiende a considerarse como varianza de error, puesto que la principal finalidad del investigador consiste en estudiar las leyes o las reglas generales que rigen nuestra conducta.

• **Análisis de datos mediante el paquete estadístico SPSS 10.0**

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Medidas repetidas** del análisis **Modelo Lineal General**.



- Indicamos el factor intra-sujetos y el número de niveles del mismo (en nuestro ejemplo, denominaremos al factor intra-sujetos «latencia» y el número de niveles es 3). A continuación escogemos la opción **Añadir** y seguidamente pulsamos la opción **Definir**.



- En el siguiente cuadro de diálogo, debemos definir los niveles del factor intra-sujetos. Para ello, basta con seleccionar cada nivel definido previamente en la matriz de datos, y pulsar el botón: 



De este modo, quedan definidos los niveles de la variable de medidas repetidas. Para obtener los resultados del análisis, debemos escoger la opción *Aceptar*.



- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

GLM

a1 a2 a3

/WSFACTOR = latencia 3 Polynomial

/METHOD = SSTYPE(3)

/CRITERIA = ALPHA(.05)

/WSDSIGN = latencia.

- Resultados:

Modelo lineal general

Factores intrasujetos

Medida: MEASURE_1

LATENCIA	Variable dependiente
1	A1
2	A2
3	A3

Contrastes multivariados^b

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Significación
LATENCIA	Traza de Pillai	0,816	8,885 ^a	2,000	4,000	0,034
	Lambda de Wilks	0,184	8,885 ^a	2,000	4,000	0,034
	Traza de Hotelling	4,443	8,885 ^a	2,000	4,000	0,034
	Raíz mayor de Roy	4,443	8,885 ^a	2,000	4,000	0,034

^a Estadístico exacto.

^b Diseño: Intercept.

Diseño intra sujetos: LATENCIA.

Prueba de esfericidad de Mauchly^b

Medida: MEASURE_1

					Epsilon ^a		
Efecto intra-sujetos	W de Mauchly	Chi-cuadrado aprox.	gl	Sig.	Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
LATENCIA	0,989	0,044	2	0,978	0,989	1,000	0,500

Contrasta la hipótesis nula de que la matriz de covarianza error de las variables dependientes transformadas es proporcional a una matriz identidad.

^a Puede usarse para corregir los grados de libertad en las pruebas de significación promediadas. Las pruebas corregidas se muestran en la tabla Pruebas de los efectos inter-sujetos.

^b Diseño: Intercept.

Diseño intra sujetos: LATENCIA.

Pruebas de efectos intrasujetos

Medida: MEASURE_1

Fuente		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
LATENCIA	Esfericidad asumida	51,444	2	25,722	11,077	0,003
	Greenhouse-Geisser	51,444	1,978	26,005	11,077	0,003
	Huynh-Feldt	51,444	2,000	25,722	11,077	0,003
	Límite-inferior	51,444	1,000	51,444	11,077	0,021
Error(LATENCIA)	Esfericidad asumida	23,222	10	2,322		
	Greenhouse-Geisser	23,222	9,891	2,348		
	Huynh-Feldt	23,222	10,000	2,322		
	Límite-inferior	23,222	5,000	4,644		

Pruebas de contrastes intrasujetos

Medida: MEASURE_1

Fuente	LATENCIA	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
LATENCIA	Lineal	36,750	1	36,750	14,412	0,013
	Cuadrático	14,694	1	14,694	7,016	0,045
Error(LATENCIA)	Lineal	12,750	5	2,550		
	Cuadrático	10,472	5	2,094		

Pruebas de los efectos inter-sujetos**Medida: MEASURE_1****Variable transformada: Promedio**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Intercept	600,889	1	600,889	466,207	0,000
Error	6,444	5	1,289		

9.1.2.3. *Análisis de la varianza mixto para el diseño factorial intrasujeto de dos factores (diseño de tratamientos $\times$ tratamientos $\times$ sujetos)*

• Modelo general de análisis

El **diseño factorial intrasujeto de dos factores**, conocido también como *diseño de tratamientos $\times$ tratamientos $\times$ sujetos*, no es sino una generalización del diseño intrasujeto simple. Se trata de una estructura de investigación en la que n sujetos seleccionados al azar de una determinada población se someten a cada una de las diferentes combinaciones de tratamientos obtenidas a partir de los valores de dos factores experimentales. A diferencia del factor sujeto, que asume un modelo de efectos aleatorios, los dos factores de tratamiento suelen asumir un modelo de efectos fijos.

El modelo matemático que subyace al análisis de la varianza bajo la hipótesis alternativa en este diseño es equivalente al del diseño de tres factores intersujetos, con la diferencia de que el tercer factor no es un factor manipulado, sino un factor sujeto. De acuerdo con el *supuesto de no aditividad de los efectos*, el modelo estructural del diseño adopta la siguiente expresión:

$$y_{ijk} = \mu + \eta_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + (\eta\alpha)_{ij} + (\eta\beta)_{ik} + (\eta\alpha\beta)_{ijk} + \varepsilon_{ijk} \quad (9.24)$$

donde los componentes adicionales:

- β_k = Efecto debido a la administración del k -ésimo nivel de la segunda variable de tratamiento.
- $(\alpha\beta)_{jk}$ = Efecto debido a la interacción entre el j -ésimo nivel de la primera variable de tratamiento y el k -ésimo nivel de la segunda variable de tratamiento.
- $(\eta\beta)_{ik}$ = Efecto debido a la interacción entre el i -ésimo sujeto y el k -ésimo nivel de la segunda variable de tratamiento.
- $(\eta\alpha\beta)_{ijk}$ = Efecto debido a la interacción entre el i -ésimo sujeto, el j -ésimo nivel de la primera variable de tratamiento y el k -ésimo nivel de la segunda variable de tratamiento.
- ε_{ijk} = Componente de error específico asociado al i -ésimo sujeto, al j -ésimo nivel de la primera variable de tratamiento y al k -ésimo nivel de la segunda variable de tratamiento.

Cabe señalar que aunque, a efectos meramente didácticos, aquí se aborde el modelo estructural no aditivo, si la interacción entre las dos variables de tratamiento no resultara estadísticamente significativa, el modelo aditivo sería mucho más adecuado para explicar los datos empíricos.

En este diseño, la interacción triple $(\eta\alpha\beta)_{ijk}$ consta de una sola observación y, por tanto, la varianza debida a tal interacción coincide con la varianza residual. A su vez, debido a las razones que se han expuesto al abordar los *supuestos básicos para el análisis en los diseños de medidas repetidas*, la presencia de un factor aleatorio en el modelo requiere asumir diferentes términos de error para contrastar las hipótesis de nulidad asociadas a los distintos parámetros de la ecuación estructural. Así, las interacciones de primer grado $(\eta\alpha)_{ij}$ y $(\eta\beta)_{ik}$ se toman como términos de error para contrastar los efectos principales de las dos variables de tratamiento (A y B), mientras que la interacción de segundo grado $(\eta\alpha\beta)_{ijk}$ se considera el término de error correspondiente a la interacción entre las dos variables manipuladas ($A \times B$). Tanto en el caso de los diseños intrasujeto simples como en el caso de los factoriales, el análisis de la varianza mixto requiere el cumplimiento de los supuestos que se han descrito en el Epígrafe 9.1.2.1 referido a los *supuestos básicos para el análisis y alternativas ante su incumplimiento*.

• Ejemplo práctico

Supongamos que en el ámbito de la psicología educativa, realizamos una investigación con el objetivo de examinar la capacidad que presentan los sujetos bilingües (inglés-castellano) para aprender diferentes tipos de contenidos incluidos en textos científicos escritos tanto en inglés como en castellano. A tal fin, se selecciona al azar una muestra de tres sujetos y se les pide que aprendan *un texto científico escrito en inglés* (a_1). Tras el aprendizaje, se registra la cantidad de *ideas principales* (b_1), de *ideas secundarias* (b_2) y de *ejemplos* (b_3) de dicho texto que recuerdan los sujetos. Transcurrido un intervalo temporal de seis meses, se les presenta *el mismo texto redactado en castellano* (a_2), pidiéndoles que lleven a cabo la misma tarea. Por último, se vuelve a medir el nivel de recuerdo de los tres tipos de contenidos arriba citados. Se parte del supuesto de que el período de seis meses es lo suficientemente amplio como para eliminar los efectos residuales que puede generar el aprendizaje del primer texto sobre el aprendizaje del segundo. En la Tabla 9.6 pueden observarse la *cantidad de unidades recordadas por el sujeto* (variable dependiente) en función del *idioma en el que se presenta el texto* (factor A) y del *tipo de contenido evaluado* (factor B).

TABLA 9.6 Matriz de datos del experimento

		A (Idioma en el que se presenta el texto)						Cuadrados de las sumas	Medias marginales
		a_1 (Inglés)			a_2 (Castellano)				
		b_1 (Princ.)	b_2 (Secun.)	b_3 (Ejem.)	b_1 (Princ.)	b_2 (Secun.)	b_3 (Ejem.)		
S (Sujetos)	S_1	19	14	8	19	25	26	12.321	18,5
	S_2	16	11	9	26	16	31	11.881	18,16
	S_3	13	20	13	25	10	17	9.604	16,33
Cuadrados de las sumas		2.304	2.025	900	4.900	2.601	5.476	101.124	
Medias marginales		16	15	10	23,3	17	24,6		17,66

La Tabla 9.6 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan las *tres categorías del factor sujeto* (S_1 , S_2 , y S_3) y en las columnas, las de los dos factores de tratamiento: en la parte superior los *dos niveles del factor A* (a_1 y a_2) y, debajo de estos, los *tres niveles del factor B* (b_1 , b_2 y b_3). Por tanto, los subíndices correspondientes a los factores S, A y B son: $i = 1, 2, 3$; $j = 1, 2$, y $k = 1, 2, 3$, respectivamente.

La Tabla 9.7 permite vislumbrar con mayor claridad la estructura subyacente al tipo de diseño que nos ocupa.

TABLA 9.7 Datos correspondientes a un diseño factorial intrasujeto de dos factores: modelo general

Factor B		Factor A														$T^2_{i..}$	Medias marg.	
		a_1					a_j					a_a						
		b_1	...	b_k	...	b_b	b_1	...	b_k	...	b_b	b_1	...	b_k	...			b_b
S U J E T O S (S)	s_1	Y_{111}	...	Y_{11k}	...	Y_{11b}	Y_{1j1}	...	Y_{1jk}	...	Y_{1jb}	Y_{1a1}	...	Y_{1ak}	...	Y_{1ab}	$T^2_{1..}$	$\bar{Y}_{1..}$
	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	s_i	Y_{i11}	...	Y_{i1k}	...	Y_{i1b}	Y_{ij1}	...	Y_{ijk}	...	Y_{ijb}	Y_{ia1}	...	Y_{iak}	...	Y_{iab}	$T^2_{i..}$	$\bar{Y}_{i..}$
	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	s_n	Y_{n11}	...	Y_{n1k}	...	Y_{n1b}	Y_{nj1}	...	Y_{njk}	...	Y_{njb}	Y_{na1}	...	Y_{nak}	...	Y_{nab}	$T^2_{n..}$	$\bar{Y}_{n..}$
$T^2_{.j.}$		$T^2_{.1.}$					$T^2_{.j.}$					$T^2_{.a.}$					$T^2_{...}$	
Medias marg.		$\bar{Y}_{.1.}$					$\bar{Y}_{.j.}$					$\bar{Y}_{.a.}$						$\bar{Y}_{...}$

Cada una de las celdillas de la Tabla 9.7 corresponde a la *puntuación obtenida por un determinado sujeto bajo una determinada combinación de tratamientos* que se le han administrado en un momento concreto. En la parte derecha (en la columna $T^2_{i..}$) se presentan los *cuadrados de las sumas correspondientes a cada sujeto* ($i = 1, 2, \dots, n$) y en la parte inferior de la tabla (en la fila $T^2_{.j.}$) los *cuadrados de las sumas correspondientes a los niveles de la primera variable de tratamiento* ($j = 1, 2, \dots, a$). A fin de que la estructura del diseño pueda percibirse con facilidad, hemos optado por no incluir en la tabla los *cuadrados de las sumas correspondientes a las categorías de la segunda variable de tratamiento* ($T^2_{.k.}$). No obstante, el lector interesado en examinar todas las dimensiones del diseño puede acceder a ellas consultando la Tabla 7.6 del Epígrafe 7.3.3. Con respecto a la notación, ya se ha señalado en varias ocasiones que la letra T hace referencia a la suma de una serie de puntuaciones y, más específicamente, a la suma de los datos que se expresan mediante los subíndices asociados a dicha letra. Por ejemplo, el símbolo $T^2_{...}$ expresa el cuadrado de la *suma de todas las puntuaciones*. Por último, en los *márgenes* se presentan las *medias aritméticas* correspondientes a cada fila o columna.

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño y, suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 9.6.

Como ya se habrá percatado el lector, la estructura del diseño que nos ocupa es similar a la del diseño factorial $A \times B \times C$ al azar (véase el Capítulo 7, Epígrafe 7.3.3).

• **Desarrollo del análisis de la varianza mixto para el diseño intrasujeto de dos factores**

Al igual que en el diseño intrasujeto simple, calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la ecuación estructural del ANOVA. Es necesario señalar que nuestras estimaciones partirán del supuesto de no aditividad de los efectos, en cuyo caso la interacción $(\eta\alpha\beta)_{ijk}$ se considera como componente residual del modelo. En concreto, desarrollaremos el ANOVA tomando como referencia la siguiente expresión:

$$Y_{ij} = \mu + \eta_i + \alpha_j + \beta_k + (\alpha\beta)_{jk} + (\eta\alpha)_{ij} + (\eta\beta)_{ik} + (\eta\alpha\beta)_{ijk} \quad (9.25)$$

Procedimiento 1

En la Tabla 9.8 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas.

TABLA 9.8 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño factorial intrasujeto de dos factores bajo el supuesto de no aditividad de los efectos

Factor de tratamiento (A)	$SCA = \left(\frac{1}{bn} \sum_j T_{.j}^2 \right) - C = \left[\frac{1}{bn} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C$	(9.26)
Factor de tratamiento (B)	$SCB = \left(\frac{1}{an} \sum_k T_{.k}^2 \right) - C = \left[\frac{1}{an} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C$	(9.27)
Factor sujeto (S)	$SCS = \left(\frac{1}{ab} \sum_i T_{i..}^2 \right) - C = \left[\frac{1}{ab} \sum_i \left(\sum_j \sum_k Y_{ijk} \right)^2 \right] - C$	(9.28)
Interacción A × S (error)	$SCAS = \frac{1}{b} \left[\sum_i \sum_j \left(\sum_k Y_{ijk} \right)^2 \right] - C - (SCA + SCS)$	(9.29)
Interacción B × S (error)	$SCBS = \frac{1}{a} \left[\sum_i \sum_k \left(\sum_j Y_{ijk} \right)^2 \right] - C - (SCB + SCS)$	(9.30)
Interacción A × B	$SCAB = \frac{1}{n} \left[\sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C - (SCA + SCB)$	(9.31)
Interacción A × B × S (error)	$SCABS = SCT - (SCA + SCB + SCS + SCAS + SCBS + SCAB)$	(9.32)
Variabilidad total	$SCT = \sum_j \sum_k \sum_i Y_{ijk}^2 - C$	(9.33)
C	$C = \frac{1}{N} T^2 = \frac{1}{abn} \left(\sum_j \sum_k \sum_i Y_{ijk} \right)^2$	(9.34)

Tras obtener el valor de C , llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{2 \cdot 3 \cdot 3} (19 + 16 + 13 + 14 + 11 + 20 + 8 + 9 + 13 + 19 + 26 + 25 + 25 + 16 + 10 + 26 + 31 + 17)^2$$

$$C = \frac{(318)^2}{18} = 5.618$$

Una vez obtenido este valor, procedemos a calcular las sumas cuadráticas de los efectos principales de los factores A , B y S , respectivamente:

$$SCA = \frac{1}{3 \cdot 3} [(19 + 16 + 13 + 14 + 11 + 20 + 8 + 9 + 13)^2 + (19 + 26 + 25 + 25 + 16 + 10 + 26 + 31 + 17)^2] - 5.618$$

$$SCA = 288$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{.j}$, a saber, a los elementos de la fila $T_{.j}^2$ de la Tabla 9.7.

$$SCB = \left[\frac{1}{2 \cdot 3} \left((19 + 16 + 13 + 19 + 26 + 25)^2 + (14 + 11 + 20 + 25 + 16 + 10)^2 + (8 + 9 + 13 + 26 + 31 + 17)^2 \right) \right] - 5.618$$

$$SCB = 41,33$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{..k}$ que, como ya se ha señalado, no han sido incluidos en la Tabla 9.7 a fin de que la representación de la estructura del diseño no resultara excesivamente compleja.

$$SCS = \frac{1}{2 \cdot 3} \left[(19 + 14 + 8 + 19 + 25 + 26)^2 + (16 + 11 + 9 + 26 + 16 + 31)^2 + (13 + 20 + 13 + 25 + 10 + 17)^2 \right] - 5.618$$

$$SCS = 16,33$$

Las puntuaciones elevadas al cuadrado corresponden a los sumatorios $T_{i..}$, es decir, a los elementos de la columna $T_{i..}^2$ de la Tabla 9.7.

Una vez calculadas las sumas de cuadrados de los tres factores que configuran el diseño, procedemos a estimar la variabilidad total y las sumas cuadráticas correspondientes a las interacciones, a saber, $SCAS$, $SCBS$, $SCAB$ y $SCABS$, respectivamente.

— *Variabilidad total:*

$$SCT = \left[(19)^2 + (16)^2 + (13)^2 + (14)^2 + (11)^2 + (20)^2 + (8)^2 + (9)^2 + (13)^2 + (19)^2 + (26)^2 + (25)^2 + (25)^2 + (16)^2 + (10)^2 + (26)^2 + (31)^2 + (17)^2 \right] - 5.618$$

$$SCT = 768$$

— SCAS (*Idioma del texto* × *Sujetos*) o residual:

$$SCAS = \frac{1}{3} \left[(19 + 14 + 8)^2 + (16 + 11 + 9)^2 + (13 + 20 + 13)^2 + \right. \\ \left. (19 + 25 + 26)^2 + (26 + 16 + 31)^2 + (25 + 10 + 17)^2 \right] - \\ - 5.618 - (288 + 16,33)$$

$$SCAS = 86,33$$

— SCBS (*Tipo de contenido* × *Sujetos*) o residual:

$$SCBS = \frac{1}{2} \left[(19 + 19)^2 + (14 + 25)^2 + (8 + 26)^2 + (16 + 26)^2 + (11 + 16)^2 + \right. \\ \left. (9 + 31)^2 + (13 + 25)^2 + (20 + 10)^2 + (13 + 17)^2 \right] - \\ - 5.618 - (41,33 + 16,33)$$

$$SCBS = 53,34$$

— SCAB (*Idioma del texto* × *Tipo de contenido*):

$$SCAB = \frac{1}{3} \left[(19 + 16 + 13)^2 + (14 + 11 + 20)^2 + (8 + 9 + 13)^2 + \right. \\ \left. (19 + 26 + 25)^2 + (25 + 16 + 10)^2 + (26 + 31 + 17)^2 \right] - \\ - 5.618 - (288 + 41,33)$$

$$SCAB = 121,33$$

— SCABS (*Idioma del texto* × *Tipo de contenido* × *Sujetos*) o residual:

$$SCABS = SCT - (SCA + SCB + SCS + SCAS + SCBS + SCAB)$$

$$SCABS = 768 - (288 + 41,33 + 16,33 + 86,33 + 53,34 + 121,34) = 161,34$$

Cabe volver a recordar que, además de calcularse mediante la aplicación de la Fórmula 9.33, la *variabilidad total* también puede obtenerse a partir de la suma de las variabilidades asociadas al resto de los efectos. Así, una vez halladas las demás sumas de cuadrados, la suma cuadrática total se calcula aplicando la siguiente expresión:

$$SCT = SCA + SCB + SCS + SCAS + SCBS + SCAB + SCABS \quad (9.35)$$

Procedimiento 2: Desarrollo mediante vectores

Omitiendo el cálculo del *vector Y* que, como es sabido, es el *vector columna* de todas las *puntuaciones directas*, comenzaremos calculando los valores correspondientes a los vectores asociados a los efectos principales de los factores A, B y S.

VECTOR B

A fin de obtener la variabilidad correspondiente a la segunda variable de tratamiento, estimaremos los parámetros asociados a dicha variable β_k . Para ello, comenzaremos calculando los promedios de las categorías del factor B.

$$\beta_k = \mu_{..k} - \mu \quad (9.39)$$

$$\mu_{..k} = \frac{1}{an} \sum_j \sum_i Y_{ijk}$$

respectivamente:

$$\mu_{..1} = \left(\frac{1}{3 \cdot 2} \right) [19 + 16 + 13 + 19 + 26 + 25] = 19,66$$

$$\mu_{..2} = \left(\frac{1}{3 \cdot 2} \right) [14 + 11 + 20 + 25 + 16 + 10] = 16$$

$$\mu_{..3} = \left(\frac{1}{3 \cdot 2} \right) [8 + 9 + 13 + 26 + 31 + 17] = 17,33$$

Por tanto:

$$\beta_1 = \mu_{..1} - \mu = 19,66 - 17,66 = 2$$

$$\beta_2 = \mu_{..2} - \mu = 16 - 17,66 = -1,66$$

$$\beta_3 = \mu_{..3} - \mu = 17,33 - 17,66 = -0,33$$

El *vector B* adopta los siguientes valores:

$$B = \{\beta\} = \begin{bmatrix} 2 \\ 2 \\ 2 \\ -1,66 \\ -1,66 \\ -1,66 \\ -0,33 \\ -0,33 \\ -0,33 \\ 2 \\ 2 \\ 2 \\ -1,66 \\ -1,66 \\ -1,66 \\ -0,33 \\ -0,33 \\ -0,33 \end{bmatrix} \quad (9.40)$$

VECTOR S

A fin de obtener la variabilidad correspondiente al factor sujeto, estimaremos los parámetros asociados a dicho factor, π_i . Para ello, comenzaremos calculando los promedios de las categorías del factor S.

$$\pi_i = \mu_{i..} - \mu \quad (9.41)$$

$$\mu_{i..} = \frac{1}{ab} \sum_j \sum_k Y_{ijk}$$

respectivamente:

$$\mu_{1..} = \frac{1}{2 \cdot 3} (19 + 14 + 8 + 19 + 25 + 26) = 18,5$$

$$\mu_{2..} = \frac{1}{2 \cdot 3} (16 + 11 + 9 + 26 + 16 + 31) = 18,16$$

$$\mu_{3..} = \frac{1}{2 \cdot 3} (13 + 20 + 13 + 25 + 10 + 17) = 16,33$$

Por tanto:

$$\pi_1 = \mu_{1..} - \mu = 18,5 - 17,66 = 0,84$$

$$\pi_2 = \mu_{2..} - \mu = 18,16 - 17,66 = 0,5$$

$$\pi_3 = \mu_{3..} - \mu = 16,33 - 17,66 = -1,33$$

El *vector S* adopta los siguientes valores:

$$S = \{\pi\} = \begin{bmatrix} 0,84 \\ 0,5 \\ -1,33 \\ 0,84 \\ 0,5 \\ -1,33 \\ 0,84 \\ 0,5 \\ -1,33 \\ 0,84 \\ 0,5 \\ -1,33 \\ 0,84 \\ 0,5 \\ -1,33 \end{bmatrix} \quad (9.42)$$

VECTOR $\{\pi\alpha\}$ O RESIDUAL

La variabilidad asociada al vector $\{\pi\alpha\}$ es la correspondiente al efecto de interacción entre los factores S y A. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\pi\alpha)_{ij}$. Para ello, comenzaremos calculando los promedios correspondientes a las combinaciones entre los niveles del factor S y del factor A.

$$\mu_{ij.} = \frac{1}{b} \sum_k Y_{ijk} \quad (9.43)$$

Aplicando la Fórmula (9.43) a los datos de nuestro ejemplo, obtenemos los siguientes valores μ_{ij} :

$$\mu_{11.} = \frac{1}{3} [19 + 14 + 8] = 13,66$$

$$\mu_{12.} = \frac{1}{3} [19 + 25 + 26] = 23,33$$

$$\mu_{21.} = \frac{1}{3} [16 + 11 + 9] = 12$$

$$\mu_{22.} = \frac{1}{3} [26 + 16 + 31] = 24,33$$

$$\mu_{31.} = \frac{1}{3} [13 + 20 + 13] = 15,33$$

$$\mu_{32.} = \frac{1}{3} [25 + 10 + 17] = 17,33$$

Una vez obtenidos tales valores, procedemos a estimar los parámetros $(\pi\alpha)_{ij}$.

$$(\pi\alpha)_{ij} = \mu_{ij.} - (\mu + \pi_i + \alpha_j)$$

$$(\pi\alpha)_{11} = \mu_{11.} - (\mu + \pi_1 + \alpha_1) = 13,66 - (17,66 + 0,84 + (-4)) = -0,84$$

$$(\pi\alpha)_{12} = \mu_{12.} - (\mu + \pi_1 + \alpha_2) = 23,33 - (17,66 + 0,84 + 4) = 0,83$$

$$(\pi\alpha)_{21} = \mu_{21.} - (\mu + \pi_2 + \alpha_1) = 12 - (17,66 + 0,5 + (-4)) = -2,16$$

$$(\pi\alpha)_{22} = \mu_{22.} - (\mu + \pi_2 + \alpha_2) = 24,33 - (17,66 + 0,5 + 4) = 2,17$$

$$(\pi\alpha)_{31} = \mu_{31.} - (\mu + \pi_3 + \alpha_1) = 15,33 - (17,66 + (-1,33) + (-4)) = 3$$

$$(\pi\alpha)_{32} = \mu_{32.} - (\mu + \pi_3 + \alpha_2) = 17,33 - (17,66 + (-1,33) + 4) = -3$$

El *vector* $\{\pi\alpha\}$ adopta los siguientes valores:

$$\{\pi\alpha\} = \begin{bmatrix} -0,84 \\ -2,16 \\ 0,83 \\ -0,84 \\ -2,16 \\ 0,83 \\ -0,84 \\ -2,16 \\ 0,83 \\ 2,17 \\ -3 \\ 0,83 \\ 2,17 \\ -3 \\ 0,83 \\ 2,17 \\ -3 \end{bmatrix} \quad (9.44)$$

VECTOR $\{\pi\beta\}$ O RESIDUAL

La variabilidad asociada al vector $\{\pi\beta\}$ es la correspondiente al efecto de interacción entre los factores S y B . Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\pi\beta)_{ik}$. Para ello, comenzaremos calculando los promedios correspondientes a las combinaciones entre los niveles del factor S y del factor B .

$$\mu_{i.k} = \frac{1}{a} \sum_j Y_{ijk} \quad (9.45)$$

Aplicando la Fórmula (9.45) a los datos de nuestro ejemplo, obtenemos los siguientes valores $\mu_{i.k}$:

$$\mu_{1.1} = \frac{1}{2} [19 + 19] = 19$$

$$\mu_{1.2} = \frac{1}{2} [14 + 25] = 19,5$$

$$\mu_{1.3} = \frac{1}{2} [8 + 26] = 17$$

$$\mu_{2.1} = \frac{1}{2} [16 + 26] = 21$$

$$\mu_{2.2} = \frac{1}{2} [11 + 16] = 13,5$$

$$\mu_{2.3} = \frac{1}{2} [9 + 31] = 20$$

$$\mu_{3.1} = \frac{1}{2} [13 + 25] = 19$$

$$\mu_{3.2} = \frac{1}{2} [20 + 10] = 15$$

$$\mu_{3.3} = \frac{1}{2} [13 + 17] = 15$$

Una vez obtenidos tales valores, procedemos a estimar los parámetros $(\pi\beta)_{ik}$.

$$(\pi\beta)_{ik} = \mu_{i.k} - (\mu + \pi_i + \beta_k)$$

$$(\pi\beta)_{11} = \mu_{1.1} - (\mu + \pi_1 + \beta_1) = 19 - (17,66 + 0,84 + 2) = -1,5$$

$$(\pi\beta)_{12} = \mu_{1.2} - (\mu + \pi_1 + \beta_2) = 19,5 - (17,66 + 0,84 + (-1,66)) = 2,66$$

$$(\pi\beta)_{13} = \mu_{1.3} - (\mu + \pi_1 + \beta_3) = 17 - (17,66 + 0,84 + (-0,33)) = -1,17$$

$$(\pi\beta)_{21} = \mu_{2.1} - (\mu + \pi_2 + \beta_1) = 21 - (17,66 + 0,5 + 2) = 0,84$$

$$(\pi\beta)_{22} = \mu_{2.2} - (\mu + \pi_2 + \beta_2) = 13,5 - (17,66 + 0,5 + (-1,66)) = -3$$

$$(\pi\beta)_{23} = \mu_{2.3} - (\mu + \pi_2 + \beta_3) = 20 - (17,66 + 0,5 + (-0,33)) = 2,17$$

$$(\pi\beta)_{31} = \mu_{3.1} - (\mu + \pi_3 + \beta_1) = 19 - (17,66 + (-1,33) + 2) = 0,67$$

$$(\pi\beta)_{32} = \mu_{3.2} - (\mu + \pi_3 + \beta_2) = 15 - (17,66 + (-1,33) + (-1,66)) = 0,33$$

$$(\pi\beta)_{33} = \mu_{3.3} - (\mu + \pi_3 + \beta_3) = 15 - (17,66 + (-1,33) + (-0,33)) = -1$$

El *vector* $\{\pi\beta\}$ adopta los siguientes valores:

$$\{\pi\beta\} = \begin{bmatrix} -1,5 \\ 0,84 \\ 0,67 \\ 2,66 \\ -3 \\ 0,33 \\ -1,17 \\ 2,17 \\ -1 \\ -1,5 \\ 0,84 \\ 0,67 \\ 2,66 \\ -3 \\ 0,33 \\ -1,17 \\ 2,17 \\ -1 \end{bmatrix} \quad (9.46)$$

VECTOR $\{\alpha\beta\}$

La variabilidad asociada al vector $\{\alpha\beta\}$ es la correspondiente al efecto de interacción entre los factores *A* y *B*. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\alpha\beta)_{jk}$. Para ello, comenzaremos calculando los promedios correspondientes a las combinaciones entre los niveles del factor *A* y del factor *B*.

$$\mu_{.jk} = \frac{1}{n} \sum_i Y_{ijk} \quad (9.47)$$

Aplicando la Fórmula (9.47) a los datos de nuestro ejemplo, obtenemos los siguientes valores $\mu_{.jk}$:

$$\mu_{.11} = \frac{1}{3} [19 + 16 + 13] = 16$$

$$\mu_{.12} = \frac{1}{3} [14 + 11 + 20] = 15$$

$$\mu_{.13} = \frac{1}{3} [8 + 9 + 13] = 10$$

$$\mu_{.21} = \frac{1}{3} [19 + 26 + 25] = 23,33$$

$$\mu_{.22} = \frac{1}{3} [25 + 16 + 10] = 17$$

$$\mu_{.23} = \frac{1}{3} [26 + 31 + 17] = 24,66$$

Una vez obtenidos tales valores, procedemos a estimar los parámetros $(\alpha\beta)_{jk}$:

$$(\alpha\beta)_{jk} = \mu_{.jk} - (\mu + \alpha_j + \beta_k) \quad (9.48)$$

$$(\alpha\beta)_{11} = \mu_{.11} - (\mu + \alpha_1 + \beta_1) = 16 - (17,66 + (-4) + 2) = 0,34$$

$$(\alpha\beta)_{12} = \mu_{.12} - (\mu + \alpha_1 + \beta_2) = 15 - (17,66 + (-4) + (-1,66)) = 3$$

$$(\alpha\beta)_{13} = \mu_{.13} - (\mu + \alpha_1 + \beta_3) = 10 - (17,66 + (-4) + (-0,33)) = -3,33$$

$$(\alpha\beta)_{21} = \mu_{.21} - (\mu + \alpha_2 + \beta_1) = 23,33 - (17,66 + 4 + 2) = -0,33$$

$$(\alpha\beta)_{22} = \mu_{.22} - (\mu + \alpha_2 + \beta_2) = 17 - (17,66 + 4 + (-1,66)) = -3$$

$$(\alpha\beta)_{23} = \mu_{.23} - (\mu + \alpha_2 + \beta_3) = 24,66 - (17,66 + 4 + (-0,33)) = 3,33$$

El vector $\{\alpha\beta\}$ adopta los siguientes valores:

$$\{\alpha\beta\} = \begin{bmatrix} 0,34 \\ 0,34 \\ 0,34 \\ 3 \\ 3 \\ 3 \\ -3,33 \\ -3,33 \\ -3,33 \\ -0,33 \\ -0,33 \\ -0,33 \\ -3 \\ -3 \\ -3 \\ 3,33 \\ 3,33 \\ 3,33 \end{bmatrix} \quad (9.49)$$

VECTOR $\{\pi\alpha\beta\}$ O RESIDUAL

La variabilidad asociada al vector $\{\pi\alpha\beta\}$ es la correspondiente al efecto de interacción entre los factores S , A y B . Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre estos tres factores, $(\pi\alpha\beta)_{ijk}$.

$$(\pi\alpha\beta)_{ijk} = Y_{ijk} - [\mu + \pi_i + \alpha_j + \beta_k + (\pi\alpha)_{ij} + (\pi\beta)_{ik} + (\alpha\beta)_{jk}] \quad (9.50)$$

Aplicando la Fórmula (9.50) a los datos de nuestro ejemplo obtenemos:

$$(\pi\alpha\beta)_{111} = 19 - (17,66 + 0,84 + (-4) + 2 + (-0,84) + (-1,5) + 0,34) = 4,5$$

$$(\pi\alpha\beta)_{112} = 14 - (17,66 + 0,84 + (-4) + (-1,66) + (-0,84) + 2,66 + 3) = -3,66$$

$$(\pi\alpha\beta)_{113} = 8 - (17,66 + 0,84 + (-4) + (-0,33) + (-0,84) + (-1,17) + (-3,33)) = -0,83$$

$$(\pi\alpha\beta)_{121} = 19 - (17,66 + 0,84 + 4 + 2 + 0,83 + (-1,5) + (-0,33)) = -4,5$$

$$(\pi\alpha\beta)_{122} = 25 - (17,66 + 0,84 + 4 + (-1,66) + 0,83 + 2,66 + (-3)) = 3,67$$

$$(\pi\alpha\beta)_{123} = 26 - (17,66 + 0,84 + 4 + (-0,33) + 0,83 + (-1,17) + 3,33) = 0,84$$

$$\begin{aligned}
(\pi\alpha\beta)_{211} &= 16 - (17,66 + 0,5 + (-4) + 2 + (-2,16) + 0,84 + 0,34) = 0,82 \\
(\pi\alpha\beta)_{212} &= 11 - (17,66 + 0,5 + (-4) + (-1,66) + (-2,16) + (-3) + 3) = 0,66 \\
(\pi\alpha\beta)_{213} &= 9 - (17,66 + 0,5 + (-4) + (-0,33) + (-2,16) + 2,17 + (-3,33)) = -1,51 \\
(\pi\alpha\beta)_{221} &= 26 - (17,66 + 0,5 + 4 + 2 + 2,17 + 0,84 + (-0,33)) = -0,84 \\
(\pi\alpha\beta)_{222} &= 16 - (17,66 + 0,5 + 4 + (-1,66) + 2,17 + (-3) + (-3)) = -0,67 \\
(\pi\alpha\beta)_{223} &= 31 - (17,66 + 0,5 + 4 + (-0,33) + 2,17 + 2,17 + 3,33) = 1,5 \\
(\pi\alpha\beta)_{311} &= 13 - (17,66 + (-1,33) + (-4) + 2 + 3 + 0,67 + 0,34) = -5,34 \\
(\pi\alpha\beta)_{312} &= 20 - (17,66 + (-1,33) + (-4) + (-1,66) + 3 + 0,33 + 3) = 3 \\
(\pi\alpha\beta)_{313} &= 13 - (17,66 + (-1,33) + (-4) + (-0,33) + 3 + (-1) + (-3,33)) = 2,33 \\
(\pi\alpha\beta)_{321} &= 25 - (17,66 + (-1,33) + 4 + 2 + (-3) + 0,67 + (-0,33)) = 5,33 \\
(\pi\alpha\beta)_{322} &= 10 - (17,66 + (-1,33) + 4 + (-1,66) + (-3) + 0,33 + (-3)) = -3 \\
(\pi\alpha\beta)_{323} &= 17 - (17,66 + (-1,33) + 4 + (-0,33) + (-3) + (-1) + 3,33) = -2,33
\end{aligned}$$

El vector $\{\pi\alpha\beta\}$ adopta los siguientes valores:

$$\{\pi\alpha\beta\} = \begin{bmatrix} 4,5 \\ -3,66 \\ -0,83 \\ -4,5 \\ 3,67 \\ 0,84 \\ 0,82 \\ 0,66 \\ -1,51 \\ -0,84 \\ -0,67 \\ 1,5 \\ -5,34 \\ 3 \\ 2,33 \\ 5,33 \\ -3 \\ -2,33 \end{bmatrix} \quad (9.51)$$

Llegados a este punto, podemos calcular la SCA, la SCB, la SCS, la SCAS, la SCBS, la SCAB, la SCABS y la SCT aplicando las Fórmulas (9.52), (9.53), (9.54), (9.55), (9.56), (9.57), (9.58) y (9.59), respectivamente, o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos.

VARIABILIDAD CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = nb \sum_j \alpha_j^2 = 3 \cdot 3 [(-4)^2 + (4)^2] = 288 \quad (9.52)$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR B (SCB):

$$SCB = na \sum_k \beta_k^2 = 3 \cdot 2 [(2)^2 + (-1,66)^2 + (-0,33)^2] = 41,18 \quad (9.53)$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR S (SCS):

$$SCS = ab \sum_i \pi_i^2 = 2 \cdot 3[(0,84)^2 + (0,5)^2 + (-1,33)^2] = 16,34 \quad (9.54)$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y S O RESIDUAL (SCAS):

$$SCAS = b \sum_j \sum_i (\pi\alpha)_{ij}^2 \quad (9.55)$$

$$SCAS = 3[(-0,84)^2 + (0,83)^2 + (-2,16)^2 + (2,17)^2 + (3)^2 + (-3)^2] = 86,3$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES B y S O RESIDUAL (SCBS):

$$SCBS = a \sum_k \sum_i (\pi\beta)_{ik}^2 \quad (9.56)$$

$$SCBS = 2 \left[\begin{array}{l} (-1,5)^2 + (2,66)^2 + (-1,17)^2 + (0,84)^2 + (-3)^2 \\ (2,17)^2 + (0,67)^2 + (0,33)^2 + (-1)^2 \end{array} \right] = 53,33$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y B (SCAB):

$$SCAB = n \sum_j \sum_k (\alpha\beta)_{jk}^2 \quad (9.57)$$

$$SCAB = 3[(0,34)^2 + (3)^2 + (-3,33)^2 + (-0,33)^2 + (-3)^2 + (3,33)^2] = 121,2$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A, B y S O RESIDUAL (SCABS):

$$SCABS = \sum_i \sum_j \sum_k (\pi\alpha\beta)_{ijk}^2 \quad (9.58)$$

$$SCABS = \left[\begin{array}{l} (4,5)^2 + (-3,66)^2 + (-0,83)^2 + (-4,5)^2 + (3,67)^2 + (0,84)^2 + \\ + (0,82)^2 + (0,66)^2 + (-1,51)^2 + (-0,84)^2 + (-0,67)^2 + (1,5)^2 + \\ + (-5,34)^2 + (3)^2 + (2,33)^2 + (5,33)^2 + (-3)^2 + (-2,33)^2 \end{array} \right] = 161,33$$

VARIABILIDAD TOTAL (SCT):

$$SCT = SCA + SCB + SCS + SCAS + SCBS + SCAB + SCABS \quad (9.59)$$

$$SCT = 288 + 41,18 + 16,34 + 86,3 + 53,33 + 121,2 + 161,33 = 767,68$$

Multiplicando cada vector por su traspuesto, obtenemos los mismos resultados:

$$SCA = \{\alpha\}^T \{\alpha\} = 288 \quad (9.60)$$

$$SCB = \{\beta\}^T \{\beta\} = 41,18 \quad (9.61)$$

$$SCS = \{\pi\}^T \{\pi\} = 16,34 \quad (9.62)$$

$$SCAS = \{\pi\alpha\}^T \{\pi\alpha\} = 86,3 \quad (9.63)$$

$$SCBS = \{\pi\beta\}^T \{\pi\beta\} = 53,33 \quad (9.64)$$

$$SCAB = \{\alpha\beta\}^T \{\alpha\beta\} = 121,2 \quad (9.65)$$

$$SCABS = \{\pi\alpha\beta\}^T \{\pi\alpha\beta\} = 161,33 \quad (9.66)$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 9.9 Análisis de la varianza mixto para un diseño factorial intrasujeto de dos factores bajo el supuesto de no aditividad de los efectos: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor de tratamiento (factor A)	SCA = 288	$a - 1 = 1$	MCA = 288	$F_A = \frac{MCA}{MCAS} = 6,67$
Factor de tratamiento (factor B)	SCB = 41,33	$b - 1 = 2$	MCB = 20,66	$F_B = \frac{MCB}{MCBS} = 1,55$
Factor sujeto (factor S)	SCS = 16,34	$n - 1 = 2$	MCS = 8,17	
Interacción A × S (error)	SCAS = 86,33	$(a - 1)(n - 1) = 2$	MCAS = 43,16	
Interacción B × S (error)	SCBS = 53,34	$(b - 1)(n - 1) = 4$	MCBS = 13,33	
Interacción A × B	SCAB = 121,33	$(a - 1)(b - 1) = 2$	MCAB = 60,66	$F_{AB} = \frac{MCAB}{MCABS} = 1,5$
Interacción A × B × S (error)	SCABS = 161,34	$(a - 1)(b - 1)(n - 1) = 4$	MCABS = 40,33	
Total	SCT = 768	$abn - 1 = 17$		

Tras la obtención de las *F* observadas, debemos determinar si la variabilidad explicada por los factores *A* y *B* y por su interacción es o no significativa. Para ello recurrimos a las *tablas de los valores críticos de la distribución F*. Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos (véase la Tabla 9.10).

Como puede apreciarse en la Tabla 9.10, el *idioma en el que se presenta el texto* (factor *A*) no ejerce una influencia estadísticamente significativa sobre su aprendizaje. De la misma forma, tampoco se observan diferencias estadísticamente significativas entre los niveles de recuerdo de los *diferentes tipos de contenidos que se evalúan tras el aprendizaje del texto*

(factor *B*). Por último, hemos de señalar que la *interacción entre ambos factores* tampoco resulta estadísticamente significativa. En definitiva, los datos de nuestro ejemplo parecen poner de manifiesto que los sujetos bilingües poseen una capacidad similar para aprender diferentes tipos de contenidos incluidos en textos científicos, tanto si éstos son presentados en inglés como si el aprendizaje se realiza en castellano.

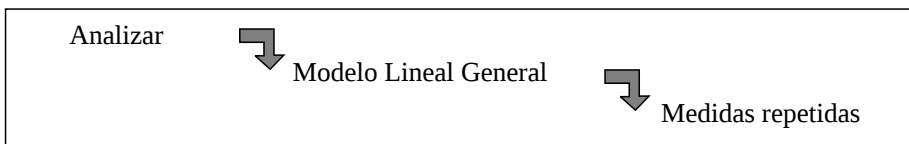
TABLA 9.10 Comparación entre las *F* observadas y las *F* teóricas con un nivel de confianza del 95 %

Fuente de variación	<i>F</i> crítica _(0,95; gl1/gl2)	<i>F</i> observada	Diferencia
Factor A	$F_{0,95; 1/2} = 18,51$	$F_A = 6,67$	$18,51 > 6,67$
Factor B	$F_{0,95; 2/4} = 6,94$	$F_B = 1,55$	$6,94 > 1,55$
Interacción A × B	$F_{0,95; 2/4} = 6,94$	$F_{AB} = 1,5$	$6,94 > 1,5$

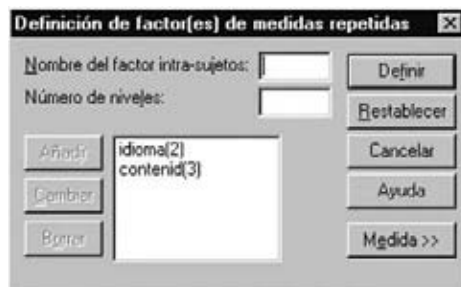
• Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

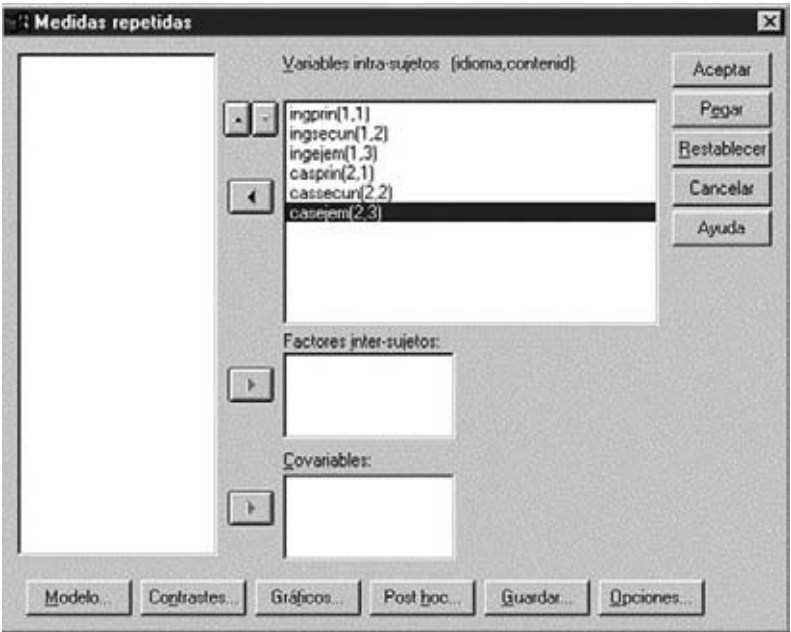
- Escogemos la opción **Medidas repetidas** del análisis **Modelo Lineal General**.



- Indicamos los factores intrasujetos y el número de niveles de los mismos. A continuación escogemos la opción **Añadir** y seguidamente pulsamos la opción **Definir**.



- En el siguiente cuadro de diálogo, debemos definir los niveles de los factores intra-sujetos, tal y como se especificó en el Epígrafe 9.1.2.2. Cabe señalar que, en el caso de los diseños factoriales, dichos niveles corresponden a los tratamientos experimentales de los que consta el diseño.



De este modo, quedan definidos los niveles de la variable de medidas repetidas. Para obtener los resultados del análisis, debemos escoger la opción *Aceptar*.

- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

```
GLM
  ingprin ingsecun ingejem casprin cassecun casejem
  /WSFACTOR = idioma 2 Polynomial contenid 3 Polynomial
  /METHOD = SSTYPE(3)
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = idioma contenid idioma*contenid.
```

- Resultados:

Modelo lineal general

Factores intrasujetos

Medida: MEASURE_1

IDIOMA	CONTENID	Variable dependiente
1	1	INGPRIN
	2	INGSECUN
	3	INGEJEM
2	1	CASPRIN
	2	CASSECUN
	3	CASEJEM

Contrastes multivariados^b

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.
IDIOMA	Traza de Pillai	0,769	6,672 ^a	1,000	2,000	0,123
	Lambda de Wilks	0,231	6,672 ^a	1,000	2,000	0,123
	Traza de Hotelling	3,336	6,672 ^a	1,000	2,000	0,123
	Raíz mayor de Roy	3,336	6,672 ^a	1,000	2,000	0,123
CONTENID	Traza de Pillai	0,862	3,127 ^a	2,000	1,000	0,371
	Lambda de Wilks	0,138	3,127 ^a	2,000	1,000	0,371
	Traza de Hotelling	6,255	3,127 ^a	2,000	1,000	0,371
	Raíz mayor de Roy	6,255	3,127 ^a	2,000	1,000	0,371
IDIOMA * CONTENID	Traza de Pillai	0,933	6,967 ^a	2,000	1,000	0,259
	Lambda de Wilks	0,067	6,967 ^a	2,000	1,000	0,259
	Traza de Hotelling	13,933	6,967 ^a	2,000	1,000	0,259
	Raíz mayor de Roy	13,933	6,967 ^a	2,000	1,000	0,259

^a Estadístico exacto.^b Diseño: Intercept.

Diseño intra sujetos: IDIOMA + CONTENID + IDIOMA * CONTENID.

Prueba de esfericidad de Mauchly^b**Medida: MEASURE_1**

					Epsilon ^a		
Efecto intra-sujetos	W de Mauchly	Chi-cuadrado aprox.	gl	Sig.	Greenhouse-Geisser	Huynh-Feldt	Límite-inferior
IDIOMA	1,000	0,000	0		1,000	1,000	1,000
CONTENID	0,263	1,337	2	0,513	0,576	0,856	0,500
IDIOMA * CONTENID	0,203	1,595	2	0,450	0,556	0,755	0,500

Contrasta la hipótesis nula de que la matriz de covarianza error de las variables dependientes transformadas es proporcional a una matriz identidad.

^a Puede usarse para corregir los grados de libertad en las pruebas de significación promediadas. Las pruebas corregidas se muestran en la tabla Pruebas de los efectos inter-sujetos.

^b Diseño: Intercept.

Diseño intra sujetos: IDIOMA + CONTENID + IDIOMA * CONTENID.

Pruebas de efectos intrasujetos**Medida: MEASURE_1**

Fuente		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
IDIOMA	Esfericidad asumida	288,000	1	288,000	6,672	0,123
	Greenhouse-Geisser	288,000	1,000	288,000	6,672	0,123
	Huynh-Feldt	288,000	1,000	288,000	6,672	0,123
	Límite-inferior	288,000	1,000	288,000	6,672	0,123

Pruebas de efectos intrasujetos (continuación)**Medida: MEASURE_1**

Fuente		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Error(IDIOMA)	Esfericidad asumida	86,333	2	43,167		
	Greenhouse-Geisser	86,333	2,000	43,167		
	Huynh-Feldt	86,333	2,000	43,167		
	Límite-inferior	86,333	2,000	43,167		
CONTENID	Esfericidad asumida	41,333	2	20,667	1,550	0,317
	Greenhouse-Geisser	41,333	1,151	35,905	1,550	0,336
	Huynh-Feldt	41,333	1,712	24,137	1,550	0,324
	Límite-inferior	41,333	1,000	41,333	1,550	0,339
Error(CONTENID)	Esfericidad asumida	53,333	4	13,333		
	Greenhouse-Geisser	53,333	2,302	23,165		
	Huynh-Feldt	53,333	3,425	15,572		
	Límite-inferior	53,333	2,000	26,667		
IDIOMA * CONTENID	Esfericidad asumida	121,333	2	60,667	1,504	0,326
	Greenhouse-Geisser	121,333	1,113	109,026	1,504	0,343
	Huynh-Feldt	121,333	1,509	80,406	1,504	0,336
	Límite-inferior	121,333	1,000	121,333	1,504	0,345
Error(IDIOMA * CONTENID)	Esfericidad asumida	161,333	4	40,333		
	Greenhouse-Geisser	161,333	2,226	72,484		
	Huynh-Feldt	161,333	3,018	53,457		
	Límite-inferior	161,333	2,000	80,667		

Pruebas de contrastes intrasujetos**Medida: MEASURE_1**

Fuente	IDIOMA	CONTENID	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
IDIOMA	Lineal		288,000	1	288,000	6,672	0,123
Error(IDIOMA)	Lineal		86,333	2	43,167		
CONTENID		Lineal	16,333	1	16,333	7,000	0,118
		Cuadrático	25,000	1	25,000	1,027	0,417
Error(CONTENID)		Lineal	4,667	2	2,333		
		Cuadrático	48,667	2	24,333		
IDIOMA * CONTENID	Lineal	Lineal	40,333	1	40,333	0,871	0,449
		Cuadrático	81,000	1	81,000	2,359	0,264
Error(IDIOMA * CONTENID)	Lineal	Lineal	92,667	2	46,333		
		Cuadrático	68,667	2	34,333		

Pruebas de los efectos intersujetos**Medida: MEASURE_1****Variable transformada: Promedio**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Intercept	5.618,000	1	5.618,000	687,918	0,001
Error	16,333	2	8,167		

9.1.2.4. Comparaciones múltiples entre medias

Como ya se ha señalado al abordar los supuestos básicos que deben tenerse en cuenta para llevar a cabo el análisis de datos cuando se trabaja con diseños intrasujeto, la *esfericidad* constituye la condición más relevante para poder calcular la prueba *F* sin generar un sesgo en su estimación (el incumplimiento de este supuesto incrementa, como mínimo, en un 0,05 el valor del α nominal). De la misma forma, el cumplimiento o incumplimiento de este supuesto resulta determinante si se pretenden realizar adecuadamente *contrastes planificados a priori entre medias específicas*. En este sentido, cabe señalar que la utilización de la media cuadrática residual común del modelo como término de error, únicamente es adecuada cuando se cumple el supuesto de *esfericidad*. En caso de que la matriz de varianzas-covarianzas del diseño no se ajuste al patrón de *esfericidad*, se recomienda la utilización de términos de error individuales para cada comparación (Keselman, 1982; Mitzel y Games, 1981). En tal circunstancia, cada contraste debe considerarse como un experimento independiente, en el que se emplea como media cuadrática residual la obtenida con el conjunto de datos que configuran dicho contraste, a saber, $MCA_{\Psi \times S}$.

Con respecto al *proceso de cálculo de las estrategias de comparaciones múltiples entre medias*, hemos de señalar que es similar al que se sigue con los diseños intersujetos, diferenciándose únicamente en la estimación de los grados de libertad del término de error. En concreto, mientras la cantidad de grados de libertad residuales que se utiliza, en el caso de los diseños intersujetos, es $N - a$, en los diseños intrasujeto dicha cantidad se obtiene mediante el producto $(a - 1)(n - 1)$, realizándose el contraste de hipótesis a partir de la siguiente expresión:

$$F = \frac{SC_{\Psi}}{MC_{A \times S}} \quad (9.67)$$

Dado que los principales procedimientos de comparaciones múltiples entre medias ya han sido desarrollados mediante ejemplos al abordar los diseños factoriales totalmente aleatorios (véase el Capítulo 7), aquí no volveremos a incidir en el proceso de cálculo. No obstante, expondremos brevemente las estrategias que resultan más adecuadas en el caso de los diseños de medidas repetidas, tanto cuando se cumple como cuando no se cumple el supuesto de *esfericidad*, y adaptaremos las fórmulas que deben aplicarse para el cálculo de los diversos *rangos críticos entre dos medias* a este tipo de diseños.

Cuando la matriz de varianzas-covarianzas del diseño es *esférica*, el **procedimiento de Tukey** permite realizar *contrastes exhaustivos a posteriori* entre todos los pares de medias

posibles, siempre y cuando tales comparaciones sean *simples*. Recordemos que cuando el diseño consta de a tratamientos, el número de comparaciones simples que se pueden llevar a cabo es igual a $a(a - 1)/2$. La fórmula para el cálculo del *rango crítico entre dos medias* en la prueba WSD de Tukey es semejante a la que se aplica en el caso de los diseños intersujetos, diferenciándose únicamente en la estimación de la cantidad de grados de libertad asociados al término de error del estadístico q . Así, la diferencia mínima entre dos medias (de los tratamientos g y h) para poder rechazar la hipótesis nula, viene determinada por la siguiente expresión:

$$|\bar{Y}_g - \bar{Y}_h| \geq \frac{q_{(\alpha, a, (a-1)(n-1))}}{\sqrt{2}} \sqrt{MC_{A \times S} \sum_{j=1}^a \frac{c_j^2}{n_j}} \quad (9.68)$$

En caso de que no se cumpla el supuesto de esfericidad y de que el investigador desee plantear *contrastes planificados a priori*, el **procedimiento de Bonferroni** permite controlar adecuadamente la tasa de error por experimento utilizando como nivel de *alpha* para cada comparación (α_{PC}), el cociente entre el nivel de *alpha* que se quiere asumir en el experimento (α_{PE}) y la cantidad de comparaciones que se realizan (c):

$$\alpha_{PC} = \frac{\alpha_{PE}}{c}$$

Como ya habrá apreciado el lector, este procedimiento se basa en la misma lógica que subyace al método aplicado en los diseños intersujetos. Por otra parte, el método de Bonferroni también resulta muy adecuado si se pretenden realizar *comparaciones a posteriori* ante la violación del supuesto de esfericidad (Maxwell y Delaney, 1990). En tal caso, dicha estrategia permite llevar a cabo todas las comparaciones posibles entre pares de medias a partir de la siguiente fórmula para el cálculo del *rango crítico entre dos medias*:

$$|\bar{Y}_g - \bar{Y}_h| \geq \sqrt{F_{(\alpha/c, 1, (a-1)(n-1))}} \sqrt{MC_{A \times S} \sum_{j=1}^a \frac{c_j^2}{n_j}} \quad (9.69)$$

No obstante, cabe señalar que en opinión de Maxwell (1980) y de Mitzel y Games (1981), la única estrategia que proporciona un adecuado control de α para todas las comparaciones dos a dos entre las a medias, ante el incumplimiento del supuesto de esfericidad, consiste en utilizar una prueba t para muestras relacionadas y un valor crítico de Dunn con los parámetros $c = a(a - 1)/2$ y g.l. = $n - 1$. En el caso de que el investigador no desee examinar todas las posibles comparaciones, c adoptaría el valor correspondiente al número de comparaciones que se fueran a realizar. Por otra parte, puede lograrse una ligera mejora en la potencia de la prueba, utilizando un valor crítico de la distribución de módulo máximo estudentizado¹ con $J^* = a(a - 1)/2$ y g.l. = $n - 1$ (Hochberg y Tamhane, 1987).

En caso de seguir esta estrategia partiendo de los datos del ejemplo práctico expuesto para el diseño intrasujeto simple (véase el Epígrafe 9.1.2.2), deberíamos utilizar pruebas t -test para realizar las comparaciones dos a dos entre los tratamientos $a_1 - a_2$, $a_1 - a_3$ y $a_2 - a_3$. Para ello, seleccionamos en el programa SPSS 10.0 la *Prueba T para muestras relacionadas*, dentro de la opción *Comparar medias* del menú *Analizar*.

¹ Los valores críticos de la distribución de módulo máximo studentizado pueden consultarse en Toothaker (1991), página 151.

Los resultados obtenidos se presentan en la siguiente tabla:

Prueba de muestras relacionadas

		Diferencias relacionadas							
		Media	Desv. típ.	Error típ. de la media	Intervalo de confianza para la diferencia		t	gl	Sig. (bilateral)
					Inferior	Superior			
Par 1	10 segundos	3,6667	2,0656	0,8433	1,4990	5,8344	4,348	5	0,007
Par 2	20 segundos								
Par 2	10 segundos	3,5000	2,2583	0,9220	1,1300	5,8700	3,796	5	0,013
	25 segundos								
Par 3	20 segundos	−0,1667	2,1370	0,8724	−2,4093	2,0760	−0,191	5	0,856
	25 segundos								

A continuación, deberíamos utilizar el valor crítico de Dunn con $c = a(a - 1)/2 = 3(3 - 1)/2 = 3$ y g.l. = $n - 1 = 6 - 1 = 5$, que para un $\alpha = 0,05$ es de 3,54 (los valores críticos de la distribución de Dunn pueden consultarse en Toothaker (1991), página 145). Como se ha indicado anteriormente, también puede adoptarse como valor crítico el de la distribución de módulo máximo studentizado, que para $J^* = 3$, g.l. = 5 y $\alpha = 0,05$ es de 3,40 (los valores críticos de la distribución de módulo máximo studentizado pueden consultarse en Toothaker (1991), página 151). Como puede percibirse el lector, este valor es algo menos conservador que el valor crítico de Dunn, lo que proporciona un ligero incremento en la potencia de la prueba. En función de estos valores críticos y de los resultados obtenidos en la prueba *t*-test, puede concluirse que existen diferencias estadísticamente significativas ($p < 0,05$) entre las condiciones a_1 (10 segundos) y a_2 (20 segundos), así como entre a_1 (10 segundos) y a_3 (25 segundos), no existiendo diferencias significativas entre los niveles de recuerdo presentados por los sujetos en los intervalos temporales de 20 (a_2) y de 25 (a_3) segundos.

Por último, si el investigador desea realizar *comparaciones post-hoc complejas* dispone del **procedimiento de Roy-Bose**, que es una extensión multivariada de la prueba de Scheffé y que, al igual que ésta, resulta válido para cualquier circunstancia, tanto si se llevan a cabo comparaciones *a priori* como *a posteriori* y tanto si tales comparaciones son *simples* o *complejas*. Este método es robusto al incumplimiento del supuesto de esfericidad, pero presenta el inconveniente de ser excesivamente conservador. En caso de aplicar el método de Roy-Bose, el *valor crítico del estadístico F* se establece mediante la siguiente expresión:

$$F_{\text{crit.}} = \frac{(n - 1)(a - 1)F_{(\alpha, (n - 1)(a - 1))}}{(n - a + 1)} \quad (9.70)$$

Para finalizar con el presente capítulo, cabe señalar que al igual que cuando utilizamos diseños unifactoriales, en el caso de rechazar la hipótesis nula respecto al efecto de interacción en el *diseño factorial* (véase el Epígrafe 7.3.5), debemos llevar a cabo contrastes específicos para determinar entre qué pares de medias existen diferencias estadísticamente significativas, tratando de controlar la tasa de error por experimento. La única diferencia en relación con el diseño unifactorial es que cuando trabajamos con un modelo de diseño que incluye un efecto de interacción estadísticamente significativo, la cantidad de posibles com-

paraciones es igual al número de tratamientos o, lo que es lo mismo, a la cantidad de combinaciones posibles entre los niveles de los factores principales, a saber, al número de *celdillas* de las que consta el diseño. En definitiva, los promedios correspondientes a las celdillas de interacción entre los factores sustituyen a los promedios de los niveles de cada factor como unidad de análisis.

9.2. DISEÑOS DE MEDIDAS PARCIALMENTE REPETIDAS: DISEÑO FACTORIAL MIXTO Y DISEÑO *SPLIT-PLOT*

9.2.1. Características generales del diseño factorial mixto y del diseño *split-plot*

El **diseño factorial mixto** es una estructura de investigación que consta de varios factores de tratamiento y en la que se combinan una o más variables de carácter intrasujeto con una o más variables de carácter intersujetos. En su *formato más simple*, el diseño factorial mixto incluye dos factores, uno de ellos responde a la estrategia de comparación intersujetos y el otro a la estrategia intrasujeto. De este modo, cada uno de los a niveles de la variable intersujetos (A) se administra aleatoriamente a diferentes grupos independientes de n sujetos y, a su vez, cada uno de los na sujetos que constituyen la muestra experimental recibe cada uno de los b niveles de una variable intrasujeto (B). En esta disposición experimental, el factor sujeto está *anidado* dentro del factor intersujetos y se *cruza* con el factor intrasujeto, pudiendo representarse el diseño mediante el símbolo $S(A) \times B$.

El **diseño *split-plot***, conocido también como **diseño de muestra o de parcela dividida**, posee una estructura similar a la del diseño factorial mixto, con la única diferencia de que las variables intersujetos no son experimentales, sino que son variables atributivas que sirven para clasificar o categorizar a los sujetos. Por tal razón, este diseño resulta muy útil en aquellas situaciones en las que los sujetos son susceptibles de ser categorizados y agrupados en función de alguna característica psicológica, clínica, biológica o social que se desea controlar o examinar. Dicha característica se conoce como *variable pronóstica*. Así, cuando el interés del investigador radica en bloquear una (o más) variable(s) atributiva(s) perturbadora(s), el diseño *split-plot* recibe el nombre de *diseño de control* y cuando, por el contrario, se pretende examinar la posible influencia diferencial ejercida por los tratamientos en función de esa(s) variable(s) atributiva(s), se denomina *diseño de interacción*.

Arnau (1995e) conceptualiza el diseño *split-plot* como una extensión del diseño de medidas repetidas de un solo grupo. De hecho, se trata de una estructura en la que los sujetos se hallan anidados en dos o más grupos que se establecen a partir de los valores que éstos presentan en características tales como el sexo, la edad, el nivel socioeconómico, etc., y en la que cada grupo recibe todos los niveles de una o más variables de tratamiento. Desde una perspectiva longitudinal, es un diseño en el que los sujetos son agrupados en función de los valores que presentan en una variable de clasificación y son observados a lo largo de una serie de puntos en el tiempo. El esquema más simple que se ajusta a esta estructura es el *diseño split-plot de dos grupos y una sola variable de respuesta*. Esta configuración puede ampliarse a situaciones más complejas o con mayor cantidad de grupos y de variables dependientes.

Aunque los diseños *split-plot* tienen su origen en la investigación agrícola, se utilizan ampliamente en el ámbito de las ciencias sociales y del comportamiento, donde a veces reciben el nombre de *diseños mixtos* o *diseños Lindquist tipo I*. Cabe señalar que en la

literatura estadística se conocen también como *diseños de perfiles* (Greenhouse y Geisser, 1959; Morrison, 1967).

Los textos de Cochran y Cox (1957), Federer (1955), Howell (1987) y Kirk (1982), entre otros, permiten obtener un conocimiento más exhaustivo acerca de los diseños de medidas parcialmente repetidas. Por tanto, remitimos al lector interesado en profundizar en tales diseños a las obras citadas.

9.2.2. El análisis de la varianza mixto para los diseños de medidas parcialmente repetidas

9.2.2.1. Supuestos básicos para el análisis de los diseños de medidas parcialmente repetidas

Los diseños de medidas parcialmente repetidas son estructuras de investigación en las que se combinan diversas características, tanto de los diseños intersujetos como de los diseños intrasujeto. Por tal razón, cuando se trabaja con diseños de medidas parcialmente repetidas, el análisis de los datos requiere el cumplimiento de los supuestos referidos a ambos tipos de diseños.

Con respecto a las *fuentes de variación intersujetos*, se deben asumir las tres condiciones básicas para llevar a cabo el análisis de la varianza con diseños de medidas independientes, a saber, el *supuesto de normalidad*, el *supuesto de homogeneidad de las varianzas* (*homocedasticidad*) y el *supuesto de independencia entre las observaciones*.

Por otra parte, las *fuentes de variación intrasujeto* requieren el cumplimiento del *supuesto de esfericidad* o *circularidad*. En lo referente a este supuesto cabe señalar que, en caso de que se incumpla, la corrección de los grados de libertad mediante la F conservadora o mediante la $\hat{\epsilon}$ de Greenhouse y Geisser, únicamente debe aplicarse sobre los factores de medidas repetidas que tengan más de dos niveles y sobre las interacciones entre cualquier factor intersujetos y aquel(los) factor(es) intrasujeto que conste(n) de más de dos niveles. Ello es así porque, como cabe comprobar fácilmente, si el factor de medidas repetidas tiene únicamente dos categorías, no se produce ningún tipo de ajuste. A su vez, es importante tener en cuenta que, dado que en los diseños de medidas parcialmente repetidas, todas las fuentes de variación intrasujeto utilizan la media cuadrática del efecto debido a la interacción entre el i -ésimo sujeto y el k -ésimo nivel de la variable intrasujeto dentro del j -ésimo nivel de la variable intersujetos ($MC_{S \times B/A}$) como término de error, un solo ajuste resulta válido para cualquiera de los efectos que se incluyen dentro de las fuentes de variación intrasujeto.

Tanto los supuestos que deben cumplirse para realizar el ANOVA con los diseños de medidas parcialmente repetidas, como las pruebas para verificar dichos supuestos y las alternativas ante su incumplimiento, ya han sido abordados en capítulos precedentes. En consecuencia, no vamos a volver a incidir en tales puntos. El lector interesado en acceder a ellos puede consultarlos en los Capítulos 6 y 9 referidos a los *diseños unifactoriales aleatorios* y a los *diseños experimentales de medidas repetidas*, respectivamente.

9.2.2.2. Modelo general de análisis

El modelo analítico que se utiliza habitualmente para llevar a cabo la **prueba de la hipótesis** con los diseños de medidas parcialmente repetidas es el *análisis de la varianza mixto*. Dejando a un lado el enfoque longitudinal o aplicado del diseño *split-plot*, en el que los

datos se analizan mediante el procedimiento que se conoce como *análisis de perfiles*, los supuestos, la descomposición de las fuentes de variación y la interpretación de los términos del modelo estructural del ANOVA son equivalentes en el diseño factorial mixto y en el diseño *split-plot*. La única diferencia entre ambos diseños es que las variables intersujetos del diseño factorial mixto son de naturaleza experimental y las del diseño *split-plot* de naturaleza atributiva. Si consideramos el *formato más simple de cualquiera de estos dos tipos de diseños*, el modelo matemático que subyace al análisis de la varianza, bajo el supuesto de la hipótesis alternativa, responde a la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \eta_{i/j} + \beta_k + (\alpha\beta)_{jk} + (\eta\beta)_{ik/j} + \varepsilon_{ijk} \quad (9.71)$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el i -ésimo sujeto bajo el j -ésimo nivel de la variable intersujetos y el k -ésimo nivel de la variable intrasujeto.

μ = Media común a todas las observaciones.

α_j = Efecto debido a la administración del j -ésimo nivel de la variable intersujetos.

$\eta_{i/j}$ = Efecto específico asociado al i -ésimo sujeto dentro del j -ésimo nivel de la variable intersujetos.

β_k = Efecto debido a la administración del k -ésimo nivel de la variable intrasujeto.

$(\alpha\beta)_{jk}$ = Efecto debido a la interacción entre el j -ésimo nivel de la variable intersujetos y el k -ésimo nivel de la variable intrasujeto.

$(\eta\beta)_{ik/j}$ = Efecto debido a la interacción entre el i -ésimo sujeto y el k -ésimo nivel de la variable intrasujeto dentro del j -ésimo nivel de la variable intersujetos.

ε_{ijk} = Componente de error específico asociado al i -ésimo sujeto, al j -ésimo nivel de la variable intersujetos y al k -ésimo nivel de la variable intrasujeto.

Se asume que los términos $\eta_{i/j}$, $(\eta\beta)_{ik/j}$ y ε_{ijk} tienen distribuciones independientes definidas por:

$$\eta_{i/j} \simeq NID(O, \sigma_{\eta}^2)$$

$$(\eta\beta)_{ik/j} \simeq NID(O, \sigma_{\eta\beta}^2)$$

$$\varepsilon_{ijk} \simeq NID(O, \sigma_{\varepsilon}^2)$$

y que se cumple el *supuesto de esfericidad o circularidad*.

Dado que la interacción de segundo orden $(\eta\beta)_{ik/j}$ coincide con la varianza residual, la ecuación estructural puede volver a formularse mediante la siguiente expresión:

$$y_{ijk} = \mu + \alpha_j + \eta_{i/j} + \beta_k + (\alpha\beta)_{jk} + (\eta\beta)_{ik/j} \quad (9.72)$$

Cabe señalar que aunque el diseño consta de tres factores, no es posible estimar todas las interacciones entre ellos, ya que la variable intersujetos no se cruza factorialmente con la variable sujeto, sino que esta última se halla anidada en la primera.

Por otra parte, debemos recordar que la variable sujeto es una variable aleatoria, lo que exige asumir diferentes términos de error para contrastar las hipótesis de nulidad asociadas a los distintos parámetros de la ecuación estructural. Así, el componente $\eta_{i/j}$ se toma como término de error para contrastar el efecto principal de la variable intersujetos (A), configurando ambos términos la *fuentes de variación intersujetos*. A su vez, la interacción $(\eta\beta)_{ik/j}$ se toma como término de error para contrastar tanto la efectividad de la variable intrasujeto (B) como el efecto de la interacción entre la variable intersujetos y la variable intrasujeto ($A \times B$). Estos tres últimos términos componen la *fuentes de variación intrasujeto*.

9.2.2.3. Ejemplo práctico: diseño split-plot

Supongamos que, en el ámbito de la psicología diferencial, realizamos una investigación para examinar si existen diferencias estadísticamente significativas en la *memoria a corto plazo* en función del *sexo*. A tal fin, se seleccionan al azar tres *mujeres* (a_1) y tres *hombres* (a_2) de la misma edad y se subdividen en dos grupos experimentales distintos. A cada uno de tales grupos se le presentan, sucesivamente y de forma aleatoria, diferentes listas de palabras, registrándose, tras cada presentación, la cantidad de *sustantivos* (b_1) y de *adjetivos* (b_2) de la lista que recuerda cada sujeto. En la Tabla 9.11 puede observarse la *cantidad de palabras recordadas por el sujeto* (variable criterio) en función de su *sexo* (factor A) y de la *categoría gramatical a la que pertenece la palabra* (factor B).

TABLA 9.11 Matriz de datos del experimento

			B (Categoría gramatical a la que pertenece la palabra)		Cuadrados de las sumas	Medias marginales
			Sujetos	b_1 (Sustantivos)		
A (Sexo del sujeto)	a_1 (Mujer)	S_1	9	11	400	10
		S_2	4	4	64	4
		S_3	6	10	256	8
	a_2 (Varón)	S_1	12	7	361	9,5
		S_2	14	12	676	13
		S_3	9	14	529	11,5
Cuadrados de las sumas			2.916	3.364	12.544	
Medias marginales			9	9,66		9,33

La Tabla 9.11 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan los *sujetos* (S_1 , S_2 y S_3) anidados en dos grupos que corresponden a las *dos categorías del factor A* (a_1 y a_2) y en las columnas se representan los *dos niveles del factor B* (b_1 y b_2). Por tanto, los subíndices del factor sujeto, del factor A y del factor B son $i = 1, 2, 3$; $j = 1, 2$ y $k = 1, 2$, respectivamente.

La Tabla 9.12 permite apreciar con mayor claridad la estructura que subyace al tipo de diseño que nos ocupa.

Cada una de las celdillas de la Tabla 9.12 corresponde a la *puntuación obtenida por un determinado sujeto bajo una determinada combinación de tratamientos* que se le han administrado en un momento concreto. En la parte derecha de la tabla (en la columna T_{ij}^2) se presentan los *cuadrados de las sumas correspondientes a cada sujeto* ($i = 1, 2, \dots, n$) *dentro de cada uno de los niveles del factor A* ($j = 1, 2, \dots, a$) y en la parte inferior (en la fila $T_{..k}^2$) los *cuadrados de las sumas correspondientes a las categorías del factor B* ($k = 1, 2, \dots, b$). Por último, en los *márgenes* se presentan las *medias aritméticas* que corresponden a cada fila o columna.

TABLA 9.12 Datos correspondientes a un diseño de medidas parcialmente repetidas de dos factores: modelo general

Factor S			Factor B			$T^2_{ij.}$	Medias marg.
			b_1	b_k	b_b		
F A C T O R A	a_1	S_{11}	Y_{111}	Y_{11k}	Y_{11b}	$T^2_{11.}$	$\bar{Y}_{11.}$
		S_{i1}	Y_{i11}	Y_{i1k}	Y_{i1b}	$T^2_{i1.}$	$\bar{Y}_{i1.}$
		S_{n1}	Y_{n11}	Y_{n1k}	Y_{n1b}	$T^2_{n1.}$	$\bar{Y}_{n1.}$
	a_j	S_{1j}	Y_{1j1}	Y_{1jk}	Y_{1jb}	$T^2_{1j.}$	$\bar{Y}_{1j.}$
		S_{ij}	Y_{ij1}	Y_{ijk}	Y_{ijb}	$T^2_{ij.}$	$\bar{Y}_{ij.}$
		S_{nj}	Y_{nj1}	Y_{njk}	Y_{njb}	$T^2_{nj.}$	$\bar{Y}_{nj.}$
	a_a	S_{1a}	Y_{1a1}	Y_{1ak}	Y_{1ab}	$T^2_{1a.}$	$\bar{Y}_{1a.}$
		S_{ia}	Y_{ia1}	Y_{iak}	Y_{iab}	$T^2_{ia.}$	$\bar{Y}_{ia.}$
		S_{na}	Y_{na1}	Y_{nak}	Y_{nab}	$T^2_{na.}$	$\bar{Y}_{na.}$
	$T^2_{..k}$			$T^2_{..1}$	$T^2_{..k}$	$T^2_{..b}$	$T^2_{...}$
Medias marg.			$\bar{Y}_{..1}$	$\bar{Y}_{..k}$	$\bar{Y}_{..b}$		$\bar{Y}_{...}$

Teniendo en cuenta la estructura general de los datos correspondiente a este tipo de diseño y, suponiendo que se cumplen todas las condiciones necesarias para aplicar el ANOVA, procederemos a su desarrollo tomando como referencia la matriz de datos de la Tabla 9.11.

9.2.2.4. Desarrollo del análisis de la varianza mixto para el diseño de medidas parcialmente repetidas

Al igual que en los diseños anteriores, calcularemos las sumas de cuadrados mediante diferentes procedimientos y, a partir de ellas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la ecuación estructural del ANOVA tomando como referencia la Fórmula (9.72).

$$y_{ijk} = \mu + \alpha_j + \eta_{i/j} + \beta_k + (\alpha\beta)_{jk} + (\eta\beta)_{ik/j} \tag{9.72}$$

Procedimiento 1

En la Tabla 9.13 presentamos las fórmulas necesarias para el cálculo de las diferentes sumas cuadráticas.

TABLA 9.13 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño de medidas parcialmente repetidas de dos factores

Variabilidad intersujetos	$SCS = \left[\frac{1}{b} \sum_i \sum_j \left(\sum_k Y_{ijk} \right)^2 \right] - C$ (9.73)
Factor intersujetos (A)	$SCA = \left[\frac{1}{bn} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C$ (9.74)
Sujetos intra A (error)	$SCS_A = SCS - SCA$ (9.75)
Factor intrasujeto (B)	$SCB = \left[\frac{1}{an} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C$ (9.76)
Interacción A × B	$SCAB = \left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C - (SCA + SCB)$ (9.77)
(Sujetos × B) intra A (error)	$SCSB_A = SCT - (SCS + SCB + SCAB)$ (9.78)
Variabilidad intrasujeto	$SC_{intra} = SCT - SCS$ (9.79)
Variabilidad total	$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$ (9.80)
C	$C = \frac{1}{abn} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2$ (9.81)

Tras obtener el valor de C , llevaremos a cabo el cálculo de las sumas de cuadrados.

$$C = \frac{1}{abn} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2 = \frac{1}{2 \cdot 2 \cdot 3} (9 + 4 + 6 + 12 + 14 + 9 + 11 + 4 + 10 + 7 + 12 + 14)^2$$

$$C = \frac{(112)^2}{12} = 1.045,33$$

Una vez obtenido este valor, procedemos a calcular las sumas cuadráticas de las diferentes fuentes de variación del modelo, a saber, la SCS , la SCA , la SCS_A , la SCB , la $SCAB$, la $SCSB_A$, la SC_{intra} y la variabilidad total o SCT .

— *Variabilidad intersujetos (SCS):*

$$SCS = \left[\frac{1}{b} \sum_i \sum_j \left(\sum_k Y_{ijk} \right)^2 \right] - C =$$

$$= \left[\frac{1}{2} [(9 + 11)^2 + (4 + 4)^2 + (6 + 10)^2 + (12 + 7)^2 + (14 + 12)^2 + (9 + 14)^2] \right] - C$$

$$SCS = \left[\frac{1}{2} [(20)^2 + (8)^2 + (16)^2 + (19)^2 + (26)^2 + (23)^2] \right] - C = 1.143 - 1.045,33 = 97,67$$

— Variabilidad correspondiente al factor A (SCA):

$$SCA = \left[\frac{1}{bn} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C =$$

$$= \left[\frac{1}{2 \cdot 3} [(9 + 4 + 6 + 11 + 4 + 10)^2 + (12 + 14 + 9 + 7 + 12 + 14)^2] \right] - C$$

$$SCA = \frac{1}{6} [(44)^2 + (68)^2] - C = 1.093,33 - 1.045,33 = 48$$

— Variabilidad residual o Sujetos intra A (SCS_A):

$$SCS_A = SCS - SCA = 97,67 - 48 = 49,67$$

— Variabilidad correspondiente al factor B (SCB):

$$SCB = \left[\frac{1}{an} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C =$$

$$= \left[\frac{1}{2 \cdot 3} [(9 + 4 + 6 + 12 + 14 + 9)^2 + (11 + 4 + 10 + 7 + 12 + 14)^2] \right] - C$$

$$SCB = \left[\frac{1}{6} (6.280) \right] - 1.045,33 = 1.046,66 - 1.045,33 = 1,33$$

— Variabilidad total (SCT):

$$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$$

$$SCT = [(9)^2 + (4)^2 + (6)^2 + (12)^2 + (14)^2 + (9)^2 + (11)^2 + (4)^2 + (10)^2 + (7)^2 + (12)^2 + (14)^2] -$$

$$- 1.045,33$$

$$SCT = 1.180 - 1.045,33 = 134,67$$

— Variabilidad correspondiente a la interacción Sexo $\times$ Categoría gramatical ($SCAB$):

$$SCAB = \left[\frac{1}{n} \sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 \right] - C - (SCA + SCB)$$

$$SCAB = \left[\frac{1}{3} [(9 + 4 + 6)^2 + (12 + 14 + 9)^2 + (11 + 4 + 10)^2 + (7 + 12 + 14)^2] \right] -$$

$$- 1.045,33 - (48 + 1,33)$$

$$SCAB = 1.100 - 1.045,33 - 49,33 = 5,34$$

— *Variabilidad residual o (Sujetos $\times$ B) intra A ($SCSB_A$):*

$$SCSB_A = SCT - (SCS + SCB + SCAB) = 134,67 - (97,67 + 1,33 + 5,34) = 30,33$$

— *Variabilidad intrasujeto (SC_{intra}):*

$$SC_{intra} = SCT - SCS = 134,67 - 97,67 = 37$$

Como ya es sabido, además de calcularse mediante la aplicación de la Fórmula (9.80), la *variabilidad total* también puede obtenerse a partir de la suma de las variabilidades asociadas al resto de los efectos. Así, una vez halladas las demás sumas de cuadrados, la suma cuadrática total se calcula aplicando la siguiente expresión:

$$SCT = SCA + SCB + SCS_A + SCAB + SCSB_A \quad (9.82)$$

A su vez, la *variabilidad intersujetos* y la *variabilidad intrasujeto* también pueden estimarse a partir de las sumas de los efectos que configuran las *fuentes de variación intersujetos e intrasujeto*, respectivamente:

$$SCS = SCA + SCS_A \quad (9.83)$$

$$SC_{intra} = SCB + SCAB + SCSB_A \quad (9.84)$$

Procedimiento 2: Desarrollo mediante vectores

En primer lugar, hallaremos los vectores asociados a cada uno de los parámetros de la ecuación estructural del ANOVA correspondiente al diseño de medidas parcialmente repetidas de dos factores (Fórmula 9.72) y, posteriormente, calcularemos las sumas de cuadrados de los diferentes efectos.

Omitiendo el cálculo del *vector Y* que, como es sabido, es el *vector columna* de todas las *puntuaciones directas*, comenzaremos calculando los valores correspondientes a los vectores asociados a los efectos principales de los factores A y B. Posteriormente, estimaremos los parámetros correspondientes a los vectores asociados a la interacción $A \times B$ y a los términos de error para las fuentes de variación intersujetos e intrasujeto, respectivamente.

VECTOR A

Calculamos la media general de la muestra y las medias correspondientes a cada una de las categorías del factor A para estimar, a partir de tales medias, los parámetros asociados al efecto del factor de tratamiento A, α_j .

$$\mu = \frac{1}{a \cdot b \cdot n} \sum_j \sum_k \sum_i Y_{ijk} \quad (9.85)$$

$$\mu = \frac{1}{2 \cdot 2 \cdot 3} (9 + 4 + 6 + 12 + 14 + 9 + 11 + 4 + 10 + 7 + 12 + 14)$$

$$\mu = \frac{112}{12} = 9,33$$

Promedios de los niveles del factor A o valores $\mu_{.j.}$:

$$\alpha_j = \mu_{.j.} - \mu \quad (9.86)$$

$$\mu_{.j.} = \frac{1}{bn} \sum_k \sum_i Y_{ijk}$$

respectivamente:

$$\mu_{.1.} = \left(\frac{1}{3 \cdot 2} \right) [9 + 4 + 6 + 11 + 4 + 10] = 7,33$$

$$\mu_{.2.} = \left(\frac{1}{3 \cdot 2} \right) [12 + 14 + 9 + 7 + 12 + 14] = 11,33$$

Por tanto:

$$\alpha_1 = \mu_{.1.} - \mu = 7,33 - 9,33 = -2$$

$$\alpha_2 = \mu_{.2.} - \mu = 11,33 - 9,33 = 2$$

El *vector A* adopta los siguientes valores:

$$A = \{\alpha\} = \begin{bmatrix} -2 \\ -2 \\ -2 \\ 2 \\ 2 \\ 2 \\ 2 \\ -2 \\ -2 \\ -2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} \quad (9.87)$$

VECTOR B

A fin de obtener la variabilidad correspondiente a la variable intrasujeto, estimaremos los parámetros asociados a dicha variable β_k . Para ello, comenzaremos calculando los promedios de las categorías del factor B .

$$\beta_k = \mu_{..k} - \mu \quad (9.88)$$

$$\mu_{..k} = \frac{1}{an} \sum_j \sum_i Y_{ijk}$$

respectivamente:

$$\mu_{..1} = \left(\frac{1}{3 \cdot 2} \right) [9 + 4 + 6 + 12 + 14 + 9] = 9$$

$$\mu_{..2} = \left(\frac{1}{3 \cdot 2} \right) [11 + 4 + 10 + 7 + 12 + 14] = 9,66$$

Por tanto:

$$\beta_1 = \mu_{..1} - \mu = 9 - 9,33 = -0,33$$

$$\beta_2 = \mu_{..2} - \mu = 9,66 - 9,33 = 0,33$$

El *vector B* toma los siguientes valores:

$$B = \{\beta\} = \begin{bmatrix} -0,33 \\ -0,33 \\ -0,33 \\ -0,33 \\ -0,33 \\ -0,33 \\ 0,33 \\ 0,33 \\ 0,33 \\ 0,33 \\ 0,33 \\ 0,33 \end{bmatrix} \quad (9.89)$$

VECTOR $\{\alpha\beta\}$

La variabilidad asociada al vector $\{\alpha\beta\}$ es la correspondiente al efecto de interacción entre los factores *A* y *B*. Calcularemos dicha variabilidad a partir de los parámetros asociados a la interacción entre ambos factores, $(\alpha\beta)_{jk}$. Para ello, comenzaremos calculando los promedios correspondientes a las combinaciones entre los niveles del factor *A* y del factor *B*.

$$\mu_{.jk} = \frac{1}{n} \sum_i Y_{ijk} \quad (9.90)$$

Aplicando la Fórmula (9.90) a los datos de nuestro ejemplo, obtenemos los siguientes valores $\mu_{.jk}$:

$$\mu_{.11} = \frac{1}{3} [9 + 4 + 6] = 6,33$$

$$\mu_{.12} = \frac{1}{3} [11 + 4 + 10] = 8,33$$

$$\mu_{.21} = \frac{1}{3} [12 + 14 + 9] = 11,66$$

$$\mu_{.22} = \frac{1}{3} [7 + 12 + 14] = 11$$

Una vez obtenidos tales valores, procedemos a estimar los parámetros $(\alpha\beta)_{jk}$:

$$(\alpha\beta)_{jk} = \mu_{.jk} - (\mu + \alpha_j + \beta_k) \quad (9.91)$$

$$(\alpha\beta)_{11} = \mu_{.11} - (\mu + \alpha_1 + \beta_1) = 6,33 - (9,33 + (-2) + (-0,33)) = -0,67$$

$$(\alpha\beta)_{12} = \mu_{.12} - (\mu + \alpha_1 + \beta_2) = 8,33 - (9,33 + (-2) + 0,33) = 0,67$$

$$(\alpha\beta)_{21} = \mu_{.21} - (\mu + \alpha_2 + \beta_1) = 11,66 - (9,33 + 2 + (-0,33)) = 0,66$$

$$(\alpha\beta)_{22} = \mu_{.22} - (\mu + \alpha_2 + \beta_2) = 11 - (9,33 + 2 + 0,33) = -0,66$$

El *vector* $\{\alpha\beta\}$ adopta los siguientes valores:

$$\{\alpha\beta\} = \begin{bmatrix} -0,67 \\ -0,67 \\ -0,67 \\ 0,66 \\ 0,66 \\ 0,66 \\ 0,67 \\ 0,67 \\ 0,67 \\ -0,66 \\ -0,66 \\ -0,66 \end{bmatrix} \quad (9.92)$$

VECTOR $\{\pi/\alpha\}$ O RESIDUAL

A fin de obtener la variabilidad correspondiente al efecto de cada uno de los sujetos dentro de cada uno de los niveles del factor A, estimaremos los parámetros asociados a dicho término de error o parámetros $(\pi/\alpha)_{ij}$. Para ello, comenzaremos calculando los promedios correspondientes a las combinaciones entre los niveles del factor sujeto y del factor A.

$$\mu_{ij.} = \frac{1}{b} \sum_k Y_{ijk} \quad (9.93)$$

Aplicando la Fórmula (9.93) a los datos de nuestro ejemplo, obtenemos los siguientes valores $\mu_{ij.}$:

$$\mu_{11.} = \frac{1}{2} [9 + 11] = 10$$

$$\mu_{12.} = \frac{1}{2} [12 + 7] = 9,5$$

$$\mu_{21.} = \frac{1}{2} [4 + 4] = 4$$

$$\mu_{22.} = \frac{1}{2} [14 + 12] = 13$$

$$\mu_{31.} = \frac{1}{2} [6 + 10] = 8$$

$$\mu_{32.} = \frac{1}{2} [9 + 14] = 11,5$$

Una vez obtenidos tales valores, procedemos a estimar los parámetros $(\pi/\alpha)_{ij}$:

$$(\pi/\alpha)_{ij} = \mu_{ij.} - (\mu + \alpha_j) \quad (9.94)$$

$$(\pi/\alpha)_{11} = \mu_{11.} - (\mu + \alpha_1) = 10 - (9,33 + (-2)) = 2,67$$

$$(\pi/\alpha)_{12} = \mu_{12.} - (\mu + \alpha_2) = 9,5 - (9,33 + 2) = -1,83$$

$$(\pi/\alpha)_{21} = \mu_{21} - (\mu + \alpha_1) = 4 - (9,33 + (-2)) = -3,33$$

$$(\pi/\alpha)_{22} = \mu_{22} - (\mu + \alpha_2) = 13 - (9,33 + 2) = 1,67$$

$$(\pi/\alpha)_{31} = \mu_{31} - (\mu + \alpha_1) = 8 - (9,33 + (-2)) = 0,67$$

$$(\pi/\alpha)_{32} = \mu_{32} - (\mu + \alpha_2) = 11,5 - (9,33 + 2) = 0,17$$

El *vector* $\{\pi/\alpha\}$ toma los siguientes valores:

$$\{\pi/\alpha\} = \begin{bmatrix} 2,67 \\ -3,33 \\ 0,67 \\ -1,83 \\ 1,67 \\ 0,17 \\ 2,67 \\ -3,33 \\ 0,67 \\ -1,83 \\ 1,67 \\ 0,17 \end{bmatrix} \quad (9.95)$$

VECTOR $\{\pi\beta/\alpha\}$ O RESIDUAL

A fin de obtener la variabilidad correspondiente al efecto de interacción entre el factor sujeto y el factor *B* dentro de cada una de las categorías del factor *A*, estimaremos los parámetros asociados a dicho término de error o parámetros $(\pi\beta/\alpha)_{ikj}$.

$$(\pi\beta/\alpha)_{ikj} = Y_{ijk} - (\mu + \alpha_j + \beta_k) + (\alpha\beta)_{jk} + (\pi/\alpha)_{ij} \quad (9.96)$$

Aplicando la Fórmula (9.96) a los datos de nuestro ejemplo, obtenemos:

$$(\pi\beta/\alpha)_{111} = 9 - (9,33 + (-2) + (-0,33) + (-0,67) + 2,67) = 0$$

$$(\pi\beta/\alpha)_{211} = 4 - (9,33 + (-2) + (-0,33) + (-0,67) + (-3,33)) = 1$$

$$(\pi\beta/\alpha)_{311} = 6 - (9,33 + (-2) + (-0,33) + (-0,67) + 0,67) = -1$$

$$(\pi\beta/\alpha)_{112} = 12 - (9,33 + 2 + (-0,33) + 0,66 + (-1,83)) = 2,17$$

$$(\pi\beta/\alpha)_{212} = 14 - (9,33 + 2 + (-0,33) + 0,66 + 1,67) = 0,67$$

$$(\pi\beta/\alpha)_{312} = 9 - (9,33 + 2 + (-0,33) + 0,66 + 0,17) = -2,83$$

$$(\pi\beta/\alpha)_{121} = 11 - (9,33 + (-2) + 0,33 + 0,67 + 2,67) = 0$$

$$(\pi\beta/\alpha)_{221} = 4 - (9,33 + (-2) + 0,33 + 0,67 + (-3,33)) = -1$$

$$(\pi\beta/\alpha)_{321} = 10 - (9,33 + (-2) + 0,33 + 0,67 + 0,67) = 1$$

$$(\pi\beta/\alpha)_{122} = 7 - (9,33 + 2 + 0,33 + (-0,66) + (-1,83)) = -2,17$$

$$(\pi\beta/\alpha)_{222} = 12 - (9,33 + 2 + 0,33 + (-0,66) + 1,67) = -0,67$$

$$(\pi\beta/\alpha)_{322} = 14 - (9,33 + 2 + 0,33 + (-0,66) + 0,17) = 2,83$$

El *vector* $\{\pi\beta/\alpha\}$ adopta los siguientes valores:

$$\{\pi\beta/\alpha\} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 2,17 \\ 0,67 \\ -2,83 \\ 0 \\ -1 \\ 1 \\ -2,17 \\ -0,67 \\ 2,83 \end{bmatrix} \quad (9.97)$$

Llegados a este punto, podemos calcular la SCA, la SCB, la SCAB, la SCS_A , la $SCSB_A$, y la SCT aplicando las Fórmulas (9.98), (9.99), (9.100), (9.101), (9.102) y (9.103), respectivamente, o bien multiplicando cada vector por su traspuesto. Procedamos a tales cálculos.

VARIABILIDAD CORRESPONDIENTE AL FACTOR A (SCA):

$$SCA = nb \sum_j \alpha_j^2 = 3 \cdot 2 [(-2)^2 + (2)^2] = 48 \quad (9.98)$$

VARIABILIDAD CORRESPONDIENTE AL FACTOR B (SCB):

$$SCB = na \sum_k \beta_k^2 = 3 \cdot 2 [(-0,33)^2 + (0,33)^2] = 1,3 \quad (9.99)$$

VARIABILIDAD CORRESPONDIENTE A LA INTERACCIÓN ENTRE LOS FACTORES A y B (SCAB):

$$SCAB = n \sum_j \sum_k (\alpha\beta)_{jk}^2 \quad (9.100)$$

$$SCAB = 3[(-0,67)^2 + (0,67)^2 + (0,66)^2 + (-0,66)^2] = 5,3$$

VARIABILIDAD RESIDUAL O SUJETOS *intra* A (SCS_A):

$$SCS_A = b \sum_j \sum_i (\pi/\alpha)_{ij}^2$$

$$SCS_A = 2[(2,67)^2 + (-3,33)^2 + (0,67)^2 + (-1,83)^2 + (1,67)^2 + (0,17)^2] \quad (9.101)$$

$$SCS_A = 49,66$$

VARIABILIDAD RESIDUAL O (SUJETOS $\times$ B) *intra* A ($SCSB_A$):

$$SCSB_A = \sum_i \sum_k \sum_j (\pi\beta/\alpha)_{ikj}^2 \quad (9.102)$$

$$SCSB_A = \left[(0)^2 + (1)^2 + (-1)^2 + (2,17)^2 + (0,67)^2 + (-2,83)^2 + (0)^2 + (-1)^2 + (1)^2 + (-2,17)^2 + (-0,67)^2 + (2,83)^2 \right] = 30,33$$

VARIABILIDAD TOTAL (SCT):

$$SCT = SCA + SCB + SCS_A + SCAB + SCSB_A \quad (9.103)$$

$$SCT = 48 + 1,33 + 49,66 + 5,34 + 30,33 = 134,66$$

Multiplicando cada vector por su traspuesto obtenemos los mismos resultados:

$$SCA = \{\alpha\}^T \{\alpha\} = 48 \quad (9.104)$$

$$SCB = \{\beta\}^T \{\beta\} = 1,33 \quad (9.105)$$

$$SCS_A = \{\pi/\alpha\}^T \{\pi/\alpha\} = 49,66 \quad (9.106)$$

$$SCAB = \{\alpha\beta\}^T \{\alpha\beta\} = 5,34 \quad (9.107)$$

$$SCSB_A = \{\beta\pi/\alpha\}^T \{\beta\pi/\alpha\} = 30,33 \quad (9.108)$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 9.14 Análisis factorial de la varianza mixto para un diseño de medidas parcialmente repetidas de dos factores: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor intersujetos (factor A)	SCA = 48	$a - 1 = 1$	$MCA = \frac{SCA}{a - 1} = 48$	$F_A = \frac{MCA}{MCS_A} = 3,86$
Factor intrasujeto (factor B)	SCB = 1,33	$b - 1 = 1$	$MCB = \frac{SCB}{b - 1} = 1,33$	$F_B = \frac{MCB}{MCSB_A} = 0,17$
Sujetos/A	$SCS_A = 49,67$	$a(n - 1) = 4$	$MCS_A = \frac{SCS_A}{a(n - 1)} = 12,42$	$F_{AB} = \frac{MCAB}{MCSB_A} = 0,70$
Interacción A × B	SCAB = 5,34	$(a - 1)(b - 1) = 1$	$MCAB = \frac{SCAB}{(a - 1)(b - 1)}$ $MCAB = 5,34$	
(Sujetos × B)/A	$SCSB_A = 30,33$	$a(n - 1)(b - 1) = 4$	$MCSB_A = \frac{SCSB_A}{a(n - 1)(b - 1)}$ $MCSB_A = 7,58$	
Total	SCT = 134,67	$abn - 1 = 11$		

Tras la obtención de las F observadas, debemos determinar si la variabilidad explicada por los factores A y B y por su interacción es o no significativa. Para ello, recurrimos a las *tablas de los valores críticos de la distribución F*. Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos:

TABLA 9.15 Comparación entre las F observadas y las F teóricas con un nivel de confianza del 95 %

Fuente de variación	F crítica _(0,95; gl1/gl2)	F observada	Diferencia
Factor A	$F_{0,95; 1/4} = 7,71$	$F_A = 3,86$	$7,71 > 3,86$
Factor B	$F_{0,95; 1/4} = 7,71$	$F_B = 0,17$	$7,71 > 0,17$
Interacción A $\times$ B	$F_{0,95; 1/4} = 7,71$	$F_{AB} = 0,70$	$7,71 > 0,70$

Como puede apreciarse en la Tabla 9.15, la variable atributiva *sexo* (*factor A*) no ejerce una influencia estadísticamente significativa sobre la cantidad de palabras recordadas por el sujeto. De la misma forma, el factor *categoría gramatical* (*factor B*) aporta al modelo una fuente de variación cuyo efecto sobre la variable criterio tampoco resulta estadísticamente significativo. Por último, hemos de señalar que no se observa un *efecto de interacción entre el factor intersujetos y el factor intrasujeto*.

9.2.2.5. Comparaciones múltiples entre medias

En el caso de los efectos principales y de las interacciones entre factores de la misma naturaleza, los procedimientos que se utilizan para llevar a cabo comparaciones múltiples entre las medias de los grupos o de los tratamientos en los diseños de medidas parcialmente repetidas son los mismos que se emplean cuando se trabaja con diseños intersujetos e intrasujeto. Como cabe deducir lógicamente, teniendo en cuenta la naturaleza de los diseños de medidas parcialmente repetidas, las comparaciones que se efectúan tras el rechazo de las hipótesis de nulidad asociadas a los efectos principales y a las interacciones que configuran la *fente de variación intersujetos* se llevan a cabo mediante procedimientos adecuados para los *diseños intersujetos*. En el caso de los efectos principales y de las interacciones que componen la *fente de variación intrasujeto*, se utilizan los métodos de comparaciones múltiples apropiados para los *diseños de medidas repetidas*. Dado que las principales estrategias de comparaciones múltiples que se emplean en las ciencias del comportamiento, tanto cuando se trabaja con diseños intersujetos (véase el Epígrafe 6.2.2.4 referido a los *diseños multigrupos aleatorios* y el Epígrafe 7.3.5 referido a los *diseños factoriales intersujetos*) como con diseños intrasujeto (véase el Epígrafe 9.1.2.4 en el que se abordan los *diseños simples y factoriales de medidas totalmente repetidas*) ya han sido expuestas anteriormente, no vamos a volver a desarrollarlas. El lector interesado en acceder a ellas puede consultarlas en los epígrafes arriba citados.

No obstante, las comparaciones múltiples entre las puntuaciones medias correspondientes a cualquier interacción entre un factor intersujetos y un factor intrasujeto son más complejas que las pruebas descritas hasta ahora. Esto se debe a dos razones: en primer lugar, a que el denominador de la prueba F general para un diseño intrasujeto y el de la prueba F global para un diseño intersujeto, son diferentes y, en segundo lugar, al supuesto de esfericidad. Como ya se ha señalado anteriormente al abordar las pruebas de comparaciones múltiples en los diseños factoriales intersujetos (véase el Epígrafe 7.3.5), las comparaciones múltiples de interés se denominan habitualmente medias de efectos simples y suelen ser de dos

tipos (Kirk, 1982): (1) diferencias entre los grupos (variable intersujetos) en cada nivel de la variable de medidas repetidas (variable intrasujeto) y (2) diferencias entre los niveles de la variable de medidas repetidas (variable intrasujeto) en cada grupo (variable intersujetos).

Las pruebas de comparaciones múltiples para el estudio de las diferencias existentes entre los grupos en cada nivel de la variable de medidas repetidas se basan en el siguiente estadístico:

$$t_{\psi_{\text{inter en intra}}} = \frac{\bar{Y}_{jk} - \bar{Y}_{j'k}}{\sqrt{(MC_{\text{intra}}/n)(2)}} \quad (9.109)$$

donde:

$$MC_{\text{intra}} = \frac{SCS_A + SCSB_A}{a(n-1) + a(n-1)(b-1)} \quad (9.110)$$

Los valores críticos, para cada comparación múltiple, pueden obtenerse mediante la siguiente expresión:

$$h_\alpha = \frac{h_1 MCS_A + h_2 MCSB_A(b-1)}{MCS_A + MCSB_A(b-1)} \quad (9.111)$$

donde,

h_α = Valor crítico para una prueba de comparaciones múltiples como, por ejemplo, la prueba HSD de Tukey.

h_1 y h_2 = Valores críticos a un nivel α para comparaciones múltiples con $glS_A = a(n-1)$ y $glSB_A = a(n-1)(b-1)$, respectivamente. Estos valores pueden extraerse a partir de los valores críticos de la distribución t (véase el Anexo A, tabla 1) o bien a partir de la distribución q de rango studentizado (véase el Anexo A, tablas 5, 6 y 7), en cuyo caso se utiliza el valor 2 como parámetro de número de medias.

Debe tenerse en cuenta que si h_1 y h_2 son valores q de rango studentizado, el valor h_α ha de dividirse entre $\sqrt{2}$ con el fin de obtener el valor crítico de t .

Volvamos a los datos de nuestro ejemplo a fin de ilustrar el cálculo de dicho estadístico (véase la Tabla 9.11). Supongamos que la prueba F global pone de manifiesto la existencia de un efecto de interacción estadísticamente significativo entre las variables sexo y categoría gramatical, y que estamos interesados en examinar la diferencia que puede existir entre varones y mujeres en cada una de las dos categorías gramaticales. Es evidente que deberíamos realizar dos pruebas de comparaciones múltiples. Consideremos, a título de ejemplo, la diferencia de medias entre varones y mujeres cuando la palabra es un sustantivo.

En primer lugar, debemos calcular el valor de MC_{intra} a partir de la Fórmula (9.110) y de los datos de la Tabla 9.14.

$$MC_{\text{intra}} = \frac{49,67 + 30,33}{4 + 4} = 10$$

El valor del estadístico t se calcula a partir de la Fórmula (9.109):

$$t_{\psi_{\text{inter en intra}}} = \frac{11,66 - 6,33}{\sqrt{(10/3)(2)}} = 2,064$$

El valor crítico para la prueba HSD de Tukey se calcula a partir de la Fórmula (9.111).

Basándonos en los valores $h_1 = 3,93$ y $h_2 = 3,93$, obtenidos a partir de la tabla de distribución de valores q de rango studentizado (véase el Anexo A, tablas 5, 6 y 7) con un parámetro de número de medias = 2, y con 4 grados de libertad, respectivamente, obtenemos el siguiente valor crítico:

$$h_x = \frac{3,93(12,42) + 3,93(7,58)(1)}{12,42 + 7,58(1)} = 3,93$$

Dado que h_1 y h_2 son valores q de rango studentizado, el valor h_x debe dividirse entre $\sqrt{2}$ con el fin de obtener el valor crítico de t .

$$t_{\text{crit.}} = \frac{h_x}{\sqrt{2}} = \frac{3,93}{\sqrt{2}} = 2,77$$

El mismo cálculo puede realizarse tomando como referencia los valores críticos $h_1 = 2,776$ y $h_2 = 2,776$ obtenidos a partir de la tabla de valores críticos de la distribución t para una prueba bilateral y con 4 grados de libertad, en ambos casos.

$$h_x = \frac{2,776(12,42) + 2,776(7,58)(1)}{12,42 + 7,58(1)} = 2,77$$

De este modo, el valor de $t = 2,064$ referido a la diferencia existente entre hombres y mujeres cuando la palabra es un sustantivo no es estadísticamente significativo, ya que no supera el valor crítico de 2,77 (recordemos que el valor de la prueba F global no resultó estadísticamente significativo, lo cual explica este resultado).

Cuando el objetivo del investigador radica en comparar las diferencias existentes entre los niveles de la variable de medidas repetidas (variable intrasujeto) en cada grupo (variable intersujetos), se utiliza la prueba t para muestras relacionadas. Este estadístico puede compararse con el valor crítico de Dunn, siendo los parámetros $gl = n - 1$ y $c = a(a - 1)/2$ si se examinan todas las posibles comparaciones dos a dos o, adoptando c el valor del número de comparaciones de interés para el investigador, en caso de que no se examinen todas las posibles comparaciones.

Retomando los datos de nuestro ejemplo (véase la Tabla 9.11), supongamos que la prueba F global pone de manifiesto la existencia de un efecto de interacción estadísticamente significativo entre las variables sexo y categoría gramatical, y que estamos interesados en examinar la diferencia que puede existir entre las dos categorías gramaticales (sustantivo y adjetivo) para cada uno de los dos sexos (varones y mujeres). Es evidente que deberíamos realizar dos pruebas de comparaciones múltiples. Consideremos, a título de ejemplo, la diferencia de medias entre las categorías sustantivo y adjetivo cuando el sujeto es varón.

Realizaremos este cálculo mediante el paquete estadístico SPSS 10.0. En primer lugar, debemos seleccionar los casos que corresponden al nivel varón del factor intersujetos sexo. A continuación, escogemos el análisis *Prueba t para muestras relacionadas* de la opción *Comparar medias* dentro del menú *Analizar*. Seguidamente, indicamos las variables relacio-

nadas (es decir, los niveles de la variable intrasujeto que se pretenden comparar), en nuestro ejemplo, sustantivo y adjetivo, y seleccionamos la opción *Aceptar*.



Los resultados obtenidos en dicho análisis se presentan a continuación:

Prueba T

Estadísticos de muestras relacionadas

		Media	N	Desviación típ.	Error típ. de la media
Par 1	SUST	11,6667	3	2,5166	1,4530
	ADJ	11,0000	3	3,6056	2,0817

Correlaciones de muestras relacionadas

		N	Correlación	Sig.
Par 1	SUST y ADJ	3	– 0,386	0,748

Prueba de muestras relacionadas

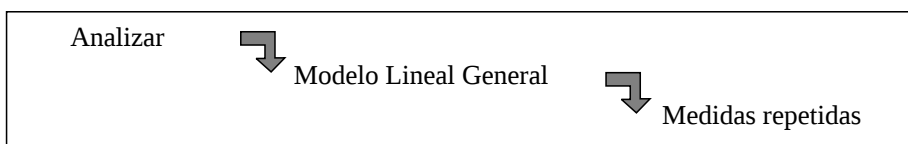
		Diferencias relacionadas							
		Media	Desv. tip.	Error tip. de la media	Intervalo de confianza para la diferencia		t	gl	Sig. (bilateral)
					Inferior	Superior			
Par 1	SUST-ADJ	0,6667	5,1316	2,9627	– 12,0809	13,4143	0,225	2	0,843

Como puede constatarse, la prueba $t(2) = 0,225$ pone de manifiesto que no existen diferencias estadísticamente significativas entre las categorías sustantivo y adjetivo, cuando los sujetos son varones (recordemos que el valor de la prueba F global no resultó estadísticamente significativo, lo cual explica este resultado).

9.2.2.6. *Análisis de datos mediante el paquete estadístico SPSS 10.0*

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Medidas repetidas** del análisis **Modelo Lineal General**.



- Indicamos el factor intra-sujetos y el número de niveles del mismo, tal y como se ha descrito en el subapartado referido al *análisis de datos mediante el paquete estadístico SPSS 10.0*, del Epígrafe 9.1.2.3.
- En el siguiente cuadro de diálogo, debemos definir los niveles de los factores intra-sujetos, tal y como se ha especificado en el Epígrafe 9.1.2.2, así como indicar el factor inter-sujetos.



De este modo quedan definidos los niveles de la variable de medidas repetidas y el factor intersujetos. Para obtener los resultados del análisis, debemos escoger la opción *Aceptar*.

- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

GLM

```
sust adj BY sexo
/WSFACTOR = categ 2 Polynomial
/METHOD = SSTYPE(3)
/CRITERIA = ALPHA(.05)
/WSDESIGN = categ
/DESIGN = sexo.
```

- Resultados:

Modelo lineal general

Factores intrasujetos

Medida: MEASURE_1

CATEG	Variable dependiente
1	SUST
2	ADJ

Factores intersujetos

		Etiqueta del valor	N
SEXO	0,00	Mujer	3
	1,00	Varón	3

Contrastes multivariados^b

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.
CATEG	Traza de Pillai	0,042	0,176 ^a	1,000	4,000	0,697
	Lambda de Wilks	0,958	0,176 ^a	1,000	4,000	0,697
	Traza de Hotelling	0,044	0,176 ^a	1,000	4,000	0,697
	Raíz mayor de Roy	0,044	0,176 ^a	1,000	4,000	0,697
CATEG * SEXO	Traza de Pillai	0,150	0,703 ^a	1,000	4,000	0,449
	Lambda de Wilks	0,850	0,703 ^a	1,000	4,000	0,449
	Traza de Hotelling	0,176	0,703 ^a	1,000	4,000	0,449
	Raíz mayor de Roy	0,176	0,703 ^a	1,000	4,000	0,449

^a Estadístico exacto.

^b Diseño: Intercept. + SEXO.

Diseño intra sujetos: CATEG.

Pruebas de efectos intrasujetos**Medida: MEASURE_1**

Fuente		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
CATEG	Esfericidad asumida	1,333	1	1,333	0,176	0,697
	Greenhouse-Geisser	1,333	1,000	1,333	0,176	0,697
	Huynh-Feldt	1,333	1,000	1,333	0,176	0,697
	Límite-inferior	1,333	1,000	1,333	0,176	0,697
CATEG * SEXO	Esfericidad asumida	5,333	1	5,333	0,703	0,449
	Greenhouse-Geisser	5,333	1,000	5,333	0,703	0,449
	Huynh-Feldt	5,333	1,000	5,333	0,703	0,449
	Límite-inferior	5,333	1,000	5,333	0,703	0,449
Error(CATEG)	Esfericidad asumida	30,333	4	7,583		
	Greenhouse-Geisser	30,333	4,000	7,583		
	Huynh-Feldt	30,333	4,000	7,583		
	Límite-inferior	30,333	4,000	7,583		

Pruebas de contrastes intrasujetos**Medida: MEASURE_1**

Fuente	CATEG	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
CATEG	Lineal	1,333	1	1,333	0,176	0,697
CATEG * SEXO	Lineal	5,333	1	5,333	0,703	0,449
Error(CATEG)	Lineal	30,333	4	7,583		

Pruebas de los efectos intersujetos**Medida: MEASURE_1****Variable transformada: Promedio**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Intercept	1.045,333	1	1.045,333	84,188	0,001
SEXO	48,000	1	48,000	3,866	0,121
Error	49,667	4	12,417		

9.3. DISEÑO *CROSS-OVER* O CONMUTATIVO Y DISEÑO DE CUADRADO LATINO INTRASUJETO

Como ya se ha señalado en el Epígrafe 9.1.1, referido a las *características generales de los diseños de medidas repetidas*, dichos diseños presentan una serie de problemas inherentes a su propia naturaleza que se conocen como *efectos de período* y *efectos residuales* o *carry-over*. No obstante, existen diseños experimentales especialmente adecuados para hacer frente a tales problemas. Entre estos diseños cabe destacar los *diseños cross-over alternativos* o *conmutativos* y los *diseños de cuadrado latino intrasujeto*.

9.3.1. Diseño *cross-over* o conmutativo

9.3.1.1. Características generales del diseño *cross-over*

Arnau (1995e) define el *diseño cross-over* como un esquema de investigación longitudinal en el que los sujetos reciben dos o más tratamientos en un determinado orden de secuenciación. Así, en el formato más simple del *diseño cross-over*, denominado *diseño* 2×2 , cada sujeto recibe dos tratamientos, pero se alterna el orden de administración de los mismos entre los dos grupos de los que consta el diseño. Así, la mitad de los sujetos (grupo 1) recibe en primer lugar el tratamiento a_1 y, tras un período de tiempo, el tratamiento a_2 . La otra mitad (grupo 2) se somete en primer lugar al tratamiento a_2 y posteriormente se le administra el tratamiento a_1 .

Como destacan Jones y Kenward (1989), los diseños conmutativos permiten sustraer de las comparaciones entre los tratamientos y entre los períodos, cualquier componente relacionado con las diferencias interindividuales. Además, la estrategia analítica asociada a este tipo de diseños posibilita estimar los efectos perturbadores derivados de la secuencia en la que se administran los tratamientos.

9.3.1.2. El análisis de la varianza para el diseño *cross-over*

• Modelo general de análisis

El procedimiento analítico que se utiliza para llevar a cabo la **prueba de la hipótesis** en los *diseños cross-over* es similar al empleado en los *diseños split-plot*. Si consideramos que el *diseño* consta de k *grupos de sujetos*, cada uno de los cuales recibe d *tratamientos en orden diferente*, a lo largo de j *períodos temporales*, el modelo estructural del diseño bajo el supuesto de la hipótesis alternativa responde a la siguiente expresión:

$$y_{ijk} = \mu + \eta_{ik} + \Pi_j + \alpha_{d[k,j]} + \lambda_{d[k,j-1]} + \varepsilon_{ijk} \quad (9.112)$$

donde:

y_{ijk} = Respuesta observada en el i -ésimo sujeto perteneciente al grupo k , en el período j .

μ = Media total del experimento.

η_{ik} = Efecto asociado al i -ésimo sujeto ($i = 1, 2, \dots, n_k$) perteneciente al grupo k ($k = 1, 2, \dots, g$).

Π_j = Efecto debido al período j ($j = 1, 2, \dots, p$).

$\alpha_{d[k,j]}$ = Efecto debido al tratamiento d ($d = 1, 2, \dots, a$) aplicado en el grupo k en el período j .

$\lambda_{d[k,j-1]}$ = Efecto residual (*carry-over*) debido al tratamiento d aplicado en el grupo k en el período $j - 1$, donde $\lambda_{[k,0]} = 0$.
 ε_{ijk} = Error aleatorio asociado al i -ésimo sujeto perteneciente al grupo k , en el período j .

Se asume que los efectos de los sujetos, η_{ik} , son independientes y siguen una distribución normal con media igual a cero y varianza común, siendo la variable sujeto una variable aleatoria. A su vez, se presupone que los errores aleatorios también son independientes y se distribuyen normalmente con una media igual a cero y una varianza igual a σ_{ε}^2 .

Como se observa en la ecuación estructural del diseño, además de los tres componentes básicos del modelo correspondiente al diseño de cuadrado latino intrasujeto que se aborda en el siguiente subapartado (Epígrafe 9.3.2), a saber, el efecto de los sujetos, el efecto del período o del orden de administración de los tratamientos y el efecto de los tratamientos, este modelo también contempla el efecto residual debido a tales tratamientos.

Con respecto a la descomposición de las fuentes de variación, cabe señalar que la *suma cuadrática total ajustada* se divide en dos grandes componentes: la *suma cuadrática intersujetos* y la *suma cuadrática intrasujeto*. Esta última se subdivide, a su vez, en tres fuentes de variación: la *suma cuadrática de los tratamientos* (ajustada a los períodos), la *suma cuadrática de los períodos* (ajustada a los tratamientos) y la *suma cuadrática residual intrasujeto*. Por último, la *suma cuadrática intersujetos* también se subdivide en dos componentes: la *suma cuadrática carry-over* y la *suma cuadrática residual intersujetos*.

• Ejemplo práctico

Supongamos que, en el ámbito de los procesos psicológicos básicos, realizamos una investigación con el objetivo de examinar la memoria a corto plazo de palabras con alto contenido emocional. A tal fin, se seleccionan aleatoriamente *dos grupos* de cinco sujetos cada uno ($n = 5$) y se les presentan dos listas de 15 palabras cada una, estando la primera de ellas formada por *palabras de alto contenido emocional* (a_1) y la segunda por *palabras neutras* (a_2). Tras una sola presentación visual de cada una de las listas, se registra la *cantidad de palabras correctamente recordadas por cada sujeto* (variable dependiente). En la Tabla 9.16 pueden observarse los resultados obtenidos en el estudio.

TABLA 9.16 Matriz de datos del experimento

Tratamiento (grupo-secuencia)	Sujeto	Período		Medias marginales
		Período 1	Período 2	
Grupo 1 (a_1a_2)	1	5	4	$\bar{Y}_{..1} = 4,4$
	2	6	4	
	3	4	2	
	4	5	3	
	5	6	5	
Grupo 2 (a_2a_1)	1	4	7	$\bar{Y}_{..2} = 4,3$
	2	3	5	
	3	3	4	
	4	5	6	
	5	2	4	
Medias marginales		$\bar{Y}_{.1.} = 4,3$	$\bar{Y}_{.2.} = 4,4$	

La Tabla 9.16 refleja la estructura correspondiente a este modelo de diseño. En las filas se representan los *sujetos* (s_1, s_2, s_3, s_4 y s_5) dentro de cada uno de los *grupos* (g_1 y g_2) de los que consta el diseño y, en las columnas, se representan los *períodos* (p_1 y p_2) a través de los cuales se les administran los *tratamientos* (a_1 y a_2) a los sujetos. Por tanto, los subíndices correspondientes a los sujetos, a los grupos, a los períodos y a los tratamientos son $i = 1, 2, 3, 4, 5$; $k = 1, 2$; $j = 1, 2$ y $d = 1, 2$, respectivamente.

La Tabla 9.17 permite vislumbrar con mayor claridad la estructura subyacente al tipo de diseño que nos ocupa.

TABLA 9.17 Tabla de datos correspondiente a un diseño *cross-over* 2×2 : modelo general

Tratamiento (grupo-secuencia)	Sujeto	Período		Medias marginales
		p_1	p_2	
g_1 ($a_1 a_2$)	S_{11} S_{i1} S_{n1}	Y_{111} Y_{i11} Y_{n11}	Y_{121} Y_{i21} Y_{n21}	$\bar{Y}_{..1}$
g_2 ($a_2 a_1$)	S_{12} S_{i2} S_{n2}	Y_{112} Y_{i12} Y_{n12}	Y_{122} Y_{i22} Y_{n22}	$\bar{Y}_{..2}$
Medias marginales		$\bar{Y}_{.1.}$	$\bar{Y}_{.2.}$	$\bar{Y}_{...}$

Cada una de las celdillas de la Tabla 9.17 corresponde a la *puntuación obtenida por un determinado sujeto en un determinado período y grupo de tratamiento*. En la parte derecha de la tabla se presentan las *medias aritméticas correspondientes a los grupos* ($k = 1, 2$) y, en la parte inferior, las correspondientes a los *períodos* ($j = 1, 2$).

Teniendo en cuenta la estructura de los datos subyacente a este modelo de diseño, procederemos a desarrollar el análisis de la varianza tomando como referencia la matriz de datos de la Tabla 9.16.

• Desarrollo del análisis de la varianza para el diseño *cross-over* 2×2

Como se ha señalado al principio del presente epígrafe (Epígrafe 9.3), el diseño *cross-over* es un diseño especialmente adecuado para hacer frente a los efectos de período y a los efectos residuales inherentes a la naturaleza de los diseños de medidas repetidas. De hecho, esa ha sido la principal razón que nos ha llevado a incluirlo en el presente texto. Sin embargo, no pertenece a la categoría de diseños que configuran el núcleo central de nuestro libro, a saber, los diseños experimentales clásicos, por lo que su tratamiento exhaustivo excede nuestros objetivos. Debido a tal circunstancia, desarrollaremos el ANOVA mediante un único procedimiento, dejando para los estudiosos de los diseños longitudinales de carácter aplicado las diferentes alternativas analíticas que se pueden utilizar con esta modalidad de diseño.

En la Tabla 9.18 presentamos las fórmulas necesarias para el cálculo de los grados de libertad y de las sumas cuadráticas de los diferentes efectos que configuran el diseño *cross-over* 2×2 .

TABLA 9.18 Fórmulas necesarias para el cálculo de los grados de libertad y de las sumas cuadráticas del ANOVA en un diseño *cross-over* 2×2

Fuentes de variación	Grados de libertad	Sumas de cuadrados
Intersujetos (<i>E-S</i>)	$n_1 + n_2 - 1$	$\sum_i \sum_k \frac{Y_{i.k}^2}{2} - \frac{Y_{...}^2}{2(n_1 + n_2)}$ (9.113)
Carry-over	1	$\frac{2n_1n_2}{(n_1 + n_2)} (\bar{Y}_{..1} - \bar{Y}_{..2})^2$ (9.114)
Residual <i>E-S</i>	$n_1 + n_2 - 2$	$\sum_i \sum_k \frac{Y_{i.k}^2}{2} - \sum_k \frac{Y_{..k}^2}{2n_k}$ (9.115)
Intrasujetos (<i>I-S</i>)	$na(p - 1)$	$\sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_k \frac{Y_{i.k}^2}{2}$ (9.116)
Tratamientos (grupos-secuencias)	1	$\frac{n_1n_2}{2(n_1 + n_2)} (\bar{Y}_{.11} - \bar{Y}_{.12} - \bar{Y}_{.21} + \bar{Y}_{.22})^2$ (9.117)
Períodos	1	$\frac{n_1n_2}{2(n_1 + n_2)} (\bar{Y}_{.11} - \bar{Y}_{.12} + \bar{Y}_{.21} - \bar{Y}_{.22})^2$ (9.118)
Residual <i>I-S</i>	$n_1 + n_2 - 2$	$\sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_k \frac{Y_{i.k}^2}{2} - \sum_j \sum_k \frac{Y_{.jk}^2}{n_k} + \sum_k \frac{Y_{..k}^2}{2n_k}$ (9.119)
Total	$2(n_1 + n_2) - 1$	$\sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{Y_{...}^2}{2(n_1 + n_2)}$ (9.120)

A fin de que el desarrollo del análisis estadístico resulte fácilmente comprensible, elaboraremos, en primer lugar, una tabla con los cálculos necesarios para obtener las variabilidades asociadas a los efectos incluidos bajo la *fuerza de variación intersujetos*. A partir de los datos de dicha tabla (véase la Tabla 9.19), estimaremos los grados de libertad y las sumas de cuadrados correspondientes a tales efectos. Posteriormente, haremos lo propio con los efectos incluidos bajo la *fuerza de variación intrasujeto*.

Tras elaborar la Tabla 9.19, procedemos a estimar los grados de libertad y las sumas de cuadrados correspondientes a los efectos que componen la fuerza de variación intersujetos.

— *Grados de libertad y variabilidad intersujetos (E-S):*

$$gl_{E-S} = n_1 + n_2 - 1 = 5 + 5 - 1 = 9$$

$$SC_{E-S} = \sum_i \sum_k \frac{Y_{i.k}^2}{2} - \frac{Y_{...}^2}{2(n_1 + n_2)} = \frac{793}{2} - \frac{87^2}{2(5 + 5)} = 18,05$$

TABLA 9.19 Cálculos necesarios para obtener las sumas cuadráticas correspondientes a los efectos que configuran la fuente de variación intersujetos

Tratamiento (grupo- secuencia)	Sujeto	Y_{ijk}		$Y_{i.k} = \sum_j Y_{ijk}$	$Y_{i.k}^2$	$Y_{..k}$	$\bar{Y}_{..k}$
		Período 1	Período 2				
Grupo 1 (a_1a_2)	1	5	4	9	81	$Y_{..1} = 44$	$\bar{Y}_{..1} = 4,4$
	2	6	4	10	100		
	3	4	2	6	36		
	4	5	3	8	64		
	5	6	5	11	121		
Grupo 2 (a_2a_1)	1	4	7	11	121	$Y_{..2} = 43$	$\bar{Y}_{..2} = 4,3$
	2	3	5	8	64		
	3	3	4	7	49		
	4	5	6	11	121		
	5	2	4	6	36		
				$Y_{...} = \sum_i \sum_j \sum_k Y_{ijk} = 87$	$\sum_i \sum_k Y_{i.k}^2 = 793$	$Y_{...} = \sum_k Y_{..k}$ $Y_{...} = 87$	$\bar{Y}_{...} = \sum_k \bar{Y}_{..k}$ $\bar{Y}_{...} = 8,7$

— Grados de libertad y variabilidad correspondientes al efecto carry-over (C-O):

$$gl_{C-O} = 1$$

$$SC_{C-O} = \frac{2n_1n_2}{(n_1 + n_2)} (\bar{Y}_{..1} - \bar{Y}_{..2})^2$$

siendo,

$$\bar{Y}_{..k} = \frac{Y_{..k}}{2n_k} = \frac{1}{2n_k} \left(\sum_i \sum_j Y_{ijk} \right)$$

$$SC_{C-O} = \frac{2 \cdot 5 \cdot 5}{(5 + 5)} (4,4 - 4,3)^2 = 0,05$$

— Grados de libertad y variabilidad correspondientes al efecto residual intersujetos (res.(E-S)):

$$gl_{res.(E-S)} = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$$

$$SC_{res.(E-S)} = \sum_i \sum_k \frac{Y_{i.k}^2}{2} - \sum_k \frac{Y_{..k}^2}{2n_k} = \frac{793}{2} - \left(\frac{44^2}{2 \cdot 5} + \frac{43^2}{2 \cdot 5} \right) = 18$$

A continuación abordamos los efectos incluidos bajo la fuente de variación intrasujeto (véase la Tabla 9.20).

TABLA 9.20 Cálculos adicionales necesarios para obtener las sumas cuadráticas correspondientes a los efectos que configuran la fuente de variación intrasujeto

		Y_{ijk}		Y_{ijk}^2		$Y_{.jk} = \sum_i Y_{ijk}$		$\bar{Y}_{.jk} = \frac{Y_{.jk}}{n_k}$	
Tratamiento (grupo- secuencia)	Sujeto	Período 1	Período 2	Período 1	Período 2	Período 1	Período 2	Período 1	Período 2
Grupo 1 (a_1a_2)	1	5	4	25	16	$Y_{.11} = 26$	$Y_{.21} = 18$	$\bar{Y}_{.11} = 5,2$	$\bar{Y}_{.21} = 3,6$
	2	6	4	36	16				
	3	4	2	16	4				
	4	5	3	25	9				
	5	6	5	36	25				
Grupo 2 (a_2a_1)	1	4	7	16	49	$Y_{.12} = 17$	$Y_{.22} = 26$	$\bar{Y}_{.12} = 3,4$	$\bar{Y}_{.22} = 5,2$
	2	3	5	9	25				
	3	3	4	9	16				
	4	5	6	25	36				
	5	2	4	4	16				
				201	212				
				$\sum_i \sum_j \sum_k Y_{ijk}^2 = 413$					

Partiendo de los cálculos llevados a cabo en las Tablas 9.19 y 9.20, procedemos a estimar los grados de libertad y las sumas de cuadrados correspondientes a los efectos que componen la fuente de variación intrasujeto.

— *Grados de libertad y variabilidad intrasujetos (I-S):*

$$gl_{I-S} = na(p - 1) = 5 \cdot 2(2 - 1) = 10$$
$$SC_{I-S} = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_k \frac{Y_{i.k}^2}{2} = 413 - \frac{793}{2} = 16,5$$

— *Grados de libertad y variabilidad correspondientes al efecto debido a los tratamientos (trat.):*

$$gl_{trat.} = 1$$
$$SC_{trat.} = \frac{n_1n_2}{2(n_1 + n_2)} (\bar{Y}_{.11} - \bar{Y}_{.12} - \bar{Y}_{.21} + \bar{Y}_{.22})^2$$

siendo,

$$\bar{Y}_{.jk} = \frac{1}{n_k} Y_{.jk} = \frac{1}{n_k} \sum_i Y_{ijk}$$
$$SC_{trat.} = \frac{5 \cdot 5}{2(5 + 5)} (5,2 - 3,4 - 3,6 + 5,2)^2 = 14,45$$

— *Grados de libertad y variabilidad correspondientes al efecto del período (per.):*

$$gl_{\text{per}} = 1$$

$$SC_{\text{per.}} = \frac{n_1 n_2}{2(n_1 + n_2)} (\bar{Y}_{.11} - \bar{Y}_{.12} + \bar{Y}_{.21} - \bar{Y}_{.22})^2 = \frac{5 \cdot 5}{2(5 + 5)} (5,2 - 3,4 + 3,6 - 5,2)^2 = 0,05$$

— *Grados de libertad y variabilidad correspondientes al efecto residual intrasujeto (res.(I-S)):*

$$gl_{\text{res.(I-S)}} = n_1 + n_2 - 2 = 5 + 5 - 2 = 8$$

$$SC_{\text{res.(I-S)}} = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_k \frac{Y_{i.k}^2}{2} - \sum_j \sum_k \frac{Y_{.jk}^2}{n_k} + \sum_k \frac{Y_{..k}^2}{2n_k}$$

$$\begin{aligned} SC_{\text{res.(I-S)}} &= 413 - \frac{793}{2} - \frac{1}{5} (26^2 + 18^2 + 17^2 + 26^2) + \frac{1}{2 \cdot 5} (44^2 + 43^2) = \\ &= 413 - 396,5 - 393 + 378,5 = 2 \end{aligned}$$

Por último, calculamos los grados de libertad y la variabilidad correspondientes a la fuente de variación total.

— *Grados de libertad y variabilidad correspondientes a la fuente de variación total:*

$$gl_{\text{total}} = 2(n_1 + n_2) - 1 = 2(5 + 5) - 1 = 19$$

$$SC_{\text{total}} = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{Y^2}{2(n_1 + n_2)} = 413 - \frac{87^2}{2(5 + 5)} = 34,55$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza (Tabla 9.21).

Tras la obtención de las F observadas, debemos determinar si la variabilidad debida al efecto residual (*carry-over*), al efecto del período y al efecto de los tratamientos es o no significativa. Para ello, recurrimos a las *tablas de los valores críticos de la distribución F*. Suponiendo que establecemos un nivel de confianza del 95 % ($\alpha = 0,05$) y que trabajamos con una hipótesis de una cola, obtendremos los siguientes valores críticos.

Como cabe apreciar en la Tabla 9.22, ni los efectos *carry-over* ni los efectos de período resultan estadísticamente significativos. Por el contrario, los tratamientos ejercen una influencia estadísticamente significativa sobre la variable criterio. En consecuencia, cabe concluir que el *contenido emocional de las palabras (factor A)* afecta significativamente a su *nivel de recuerdo*.

Para finalizar, consideramos importante destacar que cuando los efectos *carry-over* resultan significativos, la interpretación de los resultados se complica en gran medida. Ello se debe a que, como señala Arnau (1995e), no existe una única explicación para el rechazo de la hipótesis de nulidad asociada a tales efectos. De hecho, Jones y Kenward (1989) afirman que la hipótesis nula referida a los efectos *carry-over* puede rechazarse no solo porque existe un auténtico efecto *carry-over*, sino también porque dicho efecto es de carácter psicológico o debido a las expectativas de los sujetos, porque se produce una interacción entre tratamientos

y períodos, o bien porque los grupos de sujetos difieren significativamente o no son equivalentes entre sí.

TABLA 9.21 Análisis de la varianza para un diseño *cross-over* 2 × 2: ejemplo práctico

Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	<i>F</i>
Intersujetos (<i>E-S</i>)	18,05	9		
Carry-over	0,05	1	0,05	$F_{C-o} = \frac{MC_{C-o}}{MC_{res.(E-S)}} = \frac{0,05}{2,25} = 0,02$
Residual <i>E-S</i>	18	8	2,25	
Intrasujetos (<i>I-S</i>)	16,5	10		
Tratamientos (grupos-secuencias)	14,45	1	14,45	$F_{trat.} = \frac{MC_{trat.}}{MC_{res.(I-S)}} = \frac{14,45}{0,25} = 57,8$
Períodos	0,05	1	0,05	
Residual (<i>I-S</i>)	2	8	0,25	$F_{per.} = \frac{MC_{per.}}{MC_{res.(I-S)}} = \frac{0,05}{0,25} = 0,2$
Total	34,55	19		

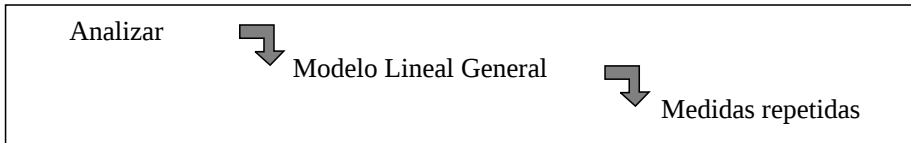
TABLA 9.22 Comparación entre las *F* observadas y las *F* teóricas con un nivel de confianza del 95 %

Fuente de variación	<i>F</i> crítica _(0,95; g11/g12)	<i>F</i> observada	Diferencia
Carry-over	$F_{0,95; 1/8} = 5,32$	$F_{C-o} = 0,02$	$5,32 > 0,02$
Períodos	$F_{0,95; 1/8} = 5,32$	$F_{per.} = 0,2$	$5,32 > 0,2$
Tratamientos (grupos-secuencias)	$F_{0,95; 1/8} = 5,32$	$F_{trat.} = 57,8$	$5,32 < 57,8$

• **Análisis de datos mediante el paquete estadístico SPSS 10.0**

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Medidas repetidas** del análisis **Modelo Lineal General**.



- Indicamos el factor intrasujetos (en nuestro ejemplo, el período) y el número de niveles del mismo, tal y como se ha descrito en el subapartado referido al *análisis de datos mediante el paquete estadístico SPSS 10.0*, del Epígrafe 9.1.2.3.
- En el siguiente cuadro de diálogo, debemos definir los niveles de los factores intrasujetos, tal y como se ha especificado en el Epígrafe 9.1.2.2, así como indicar el factor intersujetos (en nuestro ejemplo, la variable *tratamiento*).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

GLM

```

per1 per2 BY tratam
/WSFACTOR = periodo 2 Polynomial
/METHOD = SSTYPE(3)
/CRITERIA = ALPHA(.05)
/WSDESIGN = periodo
/DESIGN = tratam .
  
```

- Resultados:

A fin de que los resultados obtenidos mediante el SPSS 10.0 no induzcan a confusión, el lector debe tener en cuenta que las fuentes de variación denominadas «Tratamientos (grupos-secuencias)» y *carry-over* en la Tabla 9.21, corresponden a las fuentes de variación que en los resultados que se muestran a continuación se denominan «PERIODO * TRATAM» y «TRATAM», respectivamente.

Modelo lineal general

Factores intrasujetos

Medida: MEASURE_1

PERIODO	Variable dependiente
1	PER1
2	PER2

Factores intersujetos

		Etiqueta del valor	N
Tratamiento (grupo-secuencia)	1,00	a1-a2	5
	2,00	a2-a1	5

Contrastes multivariados^b

Efecto		Valor	F	Gl de la hipótesis	Gl del error	Sig.
PERIODO	Traza de Pillai	0,024	0,200 ^a	1,000	8,000	0,667
	Lambda de Wilks	0,976	0,200 ^a	1,000	8,000	0,667
	Traza de Hotelling	0,025	0,200 ^a	1,000	8,000	0,667
	Raíz mayor de Roy	0,025	0,200 ^a	1,000	8,000	0,667
PERIODO * TRATAM	Traza de Pillai	0,878	57,800 ^a	1,000	8,000	0,000
	Lambda de Wilks	0,122	57,800 ^a	1,000	8,000	0,000
	Traza de Hotelling	7,225	57,800 ^a	1,000	8,000	0,000
	Raíz mayor de Roy	7,225	57,800 ^a	1,000	8,000	0,000

^a Estadístico exacto.

^b Diseño: Intercept. + TRATAM.

Diseño intra sujetos: PERIODO.

Pruebas de efectos intrasujetos

Medida: MEASURE_1

Fuente		Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
PERIODO	Esfericidad asumida	5,000E-02	1	5,000E-02	0,200	0,667
	Greenhouse-Geisser	5,000E-02	1,000	5,000E-02	0,200	0,667
	Huynh-Feldt	5,000E-02	1,000	5,000E-02	0,200	0,667
	Límite-inferior	5,000E-02	1,000	5,000E-02	0,200	0,667
PERIODO * TRATAM	Esfericidad asumida	14,450	1	14,450	57,800	0,000
	Greenhouse-Geisser	14,450	1,000	14,450	57,800	0,000
	Huynh-Feldt	14,450	1,000	14,450	57,800	0,000
	Límite-inferior	14,450	1,000	14,450	57,800	0,000
Error(PERIODO)	Esfericidad asumida	2,000	8	0,250		
	Greenhouse-Geisser	2,000	8,000	0,250		
	Huynh-Feldt	2,000	8,000	0,250		
	Límite-inferior	2,000	8,000	0,250		

Pruebas de contrastes intrasujetos

Medida: MEASURE_1

Fuente	PERIODO	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
PERIODO	Lineal	5,000E-02	1	5,000E-02	0,200	0,667
PERIODO * TRATAM	Lineal	14,450	1	14,450	57,800	0,000
Error(PERIODO)	Lineal	2,000	8	0,250		

Pruebas de los efectos intersujetos**Medida: MEASURE_1****Variable transformada: Promedio**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.
Intercept	378,450	1	378,450	168,200	0,000
TRATAM	5,000E-02	1	5,000E-02	0,022	0,885
Error	18,000	8	2,250		

9.3.2. Diseño de cuadrado latino intrasujeto*9.3.2.1 Características generales del diseño de cuadrado latino intrasujeto*

El *diseño de cuadrado latino intrasujeto* es un esquema de investigación en el que los sujetos reciben diferentes tratamientos siguiendo la disposición que se conoce como *cuadrado latino*. En tal disposición, las filas corresponden a la variable sujeto y las columnas al orden en el que se administran los tratamientos. Tanto estas dos dimensiones de variación como la dimensión correspondiente a los tratamientos adoptan los mismos valores. Así, por ejemplo, el *diseño intrasujeto de cuadrado latino* 4×4 se representa mediante la siguiente matriz de doble entrada:

TABLA 9.23 Representación del diseño intrasujeto de cuadrado latino 4×4

		Período u orden			
		O_1	O_2	O_3	O_4
Sujetos	S_1	A_1	A_2	A_3	A_4
	S_2	A_2	A_3	A_4	A_1
	S_3	A_3	A_4	A_1	A_2
	S_4	A_4	A_1	A_2	A_3

Como se observa en esta matriz (Tabla 9.23), cada sujeto recibe la secuencia de los tratamientos en diferente orden, y los distintos órdenes de administración de los tratamientos se *contrabalancean* a través de los sujetos. De esta manera, cualquiera de los tratamientos ocupa el mismo lugar tantas veces como los restantes, a lo largo de los distintos sujetos. En consecuencia, ningún tratamiento presenta ventajas o desventajas derivadas de ocupar una determinada posición. Así, la estructura interna del cuadrado latino permite controlar los efectos de período mediante la estrategia de *contrabalanceo*. Además, la técnica analítica asociada a este tipo de diseños posibilita estimar de forma independiente el efecto debido a los tratamientos, el debido a los sujetos y el derivado del orden en el que se administran los

tratamientos. No obstante, la principal desventaja del diseño de cuadrado latino intrasujeto radica en que no permite controlar los efectos residuales o *carry-over*. En consecuencia, cuando se utiliza este tipo de diseños, resulta muy recomendable establecer intervalos temporales amplios entre los diferentes tratamientos a fin de eliminar dichos efectos.

9.3.2.2. El análisis de la varianza para el diseño de cuadrado latino intrasujeto

• Modelo general de análisis

El procedimiento analítico que se les aplica habitualmente a los datos generados a partir de los diseños intrasujeto de cuadrado latino es el *análisis de la varianza*. El modelo matemático que subyace a dicho análisis, bajo el supuesto de la hipótesis alternativa, responde a la siguiente expresión:

$$y_{ijk} = \mu + \eta_i + \alpha_j + \Pi_k + \varepsilon_{ijk} \quad (9.121)$$

donde:

y_{ijk} = Puntuación obtenida en la variable dependiente por el i -ésimo sujeto bajo el j -ésimo tratamiento administrado en el k -ésimo orden.

μ = Media global de las puntuaciones de los sujetos.

η_i = Efecto asociado al i -ésimo sujeto.

α_j = Efecto debido a la administración del j -ésimo tratamiento.

Π_k = Efecto asociado al k -ésimo orden o período.

ε_{ijk} = Componente de error aleatorio del modelo.

Se asume que los efectos debidos al orden, a los sujetos y a los tratamientos son aditivos y que, por tanto, las interacciones son nulas. A su vez, se presupone que se cumple el supuesto de homogeneidad de las covarianzas y que los términos η_i y ε_{ijk} tienen distribuciones independientes definidas por:

$$\eta_i \simeq NID(O, \sigma_\eta^2)$$

$$\varepsilon_{ijk} \simeq NID(O, \sigma_\varepsilon^2)$$

En relación con tales supuestos, Arnau (1986) afirma que la presencia de una interacción entre la variable de tratamiento y la de orden constituye un serio problema en este tipo de diseños. Ello se debe a que, en caso de producirse, la varianza de dicha interacción formaría parte de la varianza de error, dando lugar a un sesgo negativo en la estimación de la F y a una disminución en la potencia del diseño. Por tal razón, Arnau recomienda aplicar la *prueba de no aditividad de Tukey* (1949) a fin de contrastar la hipótesis de nulidad asociada a los efectos de la interacción (el lector interesado en el desarrollo matemático de dicha prueba puede consultarla en el Epígrafe 8.1 referido a los *diseños de bloques aleatorios*).

En lo que respecta a la descomposición de las fuentes de variación del diseño, cabe señalar que la *suma de cuadrados total* incluye cuatro componentes: la *suma cuadrática asociada a los sujetos* o a las filas, la *suma cuadrática asociada al orden de administración de los tratamientos* o a las columnas, la *suma cuadrática debida a los tratamientos* y la *suma cuadrática residual*. Esta última fuente de variación se toma básicamente como término de

error para contrastar el efecto debido a los tratamientos, aunque también puede utilizarse para contrastar los efectos asociados a los sujetos y al orden en el que se administran los tratamientos.

• Ejemplo práctico

Supongamos que en el ámbito de los procesos psicológicos básicos, realizamos una investigación con el objetivo de examinar la memoria a corto plazo. En concreto, se pretende analizar la influencia que ejerce el *nivel de familiaridad de una serie de palabras presentadas acústicamente* (factor A) sobre su posterior *recuerdo* (variable dependiente). A tal fin, se seleccionan palabras pertenecientes a tres categorías distintas respecto a su nivel de familiaridad, a saber: (a_1) *palabras de bajo nivel de familiaridad*, (a_2) *palabras de nivel de familiaridad medio* y (a_3) *palabras de nivel de familiaridad alto*. Tras escoger al azar una muestra de tres sujetos, se los somete a ocho ensayos sucesivos consistentes en la presentación de secuencias de cinco palabras pertenecientes a las distintas condiciones experimentales. Con el objetivo de controlar los efectos de período, se utiliza un modelo de diseño intrasujeto de cuadrado latino 3×3 , en el que los tratamientos se administran atendiendo a la siguiente configuración o matriz de doble entrada:

TABLA 9.24 Representación del diseño intrasujeto de cuadrado latino 3×3 : ejemplo práctico

		Período u orden		
		O_1	O_2	O_3
Sujetos	S_1	a_1	a_2	a_3
	S_2	a_2	a_3	a_1
	S_3	a_3	a_1	a_2

En la Tabla 9.25 se puede observar la cantidad de palabras correctamente recordadas por los sujetos, bajo cada una de las condiciones de tratamiento.

Al igual que en el caso del diseño de cuadrado latino intersujetos (véase el Epígrafe 8.2), el principal objetivo que se persigue al abordar la modalidad de diseño que aquí nos ocupa, es el de ilustrar la técnica subyacente a dicha estructura de investigación. Por ello, desarrollaremos el ANOVA utilizando únicamente el primero de los procedimientos que hemos empleado para calcular las sumas de cuadrados en la mayoría de los diseños abordados con anterioridad. Tras obtener las sumas cuadráticas, estimaremos las varianzas y las razones F para cada uno de los parámetros de la ecuación estructural del diseño (véase la Fórmula 9.121). Cabe señalar que el desarrollo del ANOVA mediante vectores no plantea ninguna dificultad y que se realiza exactamente igual que en el caso de los diseños que hemos presentado en los epígrafes precedentes. Dicho esto, procedamos a desarrollar el análisis de la varianza para el diseño intrasujeto de cuadrado latino 3×3 , tomando como referencia la matriz de datos de la Tabla 9.25.

TABLA 9.25 Matriz de datos del experimento

		Período u orden			Sumatorios y medias marginales	A (Nivel de familiaridad de las palabras)
		O ₁	O ₂	O ₃		
Sujetos	S ₁	a ₁ (bajo) 7	a ₂ (medio) 5	a ₃ (alto) 4	Σ Y _{1..} = 16 Ȳ _{1..} = 5,33	Σ Y _{.1.} = 19 Ȳ _{.1.} = 6,33
	S ₂	a ₂ (medio) 5	a ₃ (alto) 4	a ₁ (bajo) 6	Σ Y _{2..} = 15 Ȳ _{2..} = 5	Σ Y _{.2.} = 14 Ȳ _{.2.} = 4,66
	S ₃	a ₃ (alto) 3	a ₁ (bajo) 6	a ₂ (medio) 4	Σ Y _{3..} = 13 Ȳ _{3..} = 4,33	Σ Y _{.3.} = 11 Ȳ _{.3.} = 3,66
Sumatorios y medias marginales		Σ Y _{..1} = 15 Ȳ _{..1} = 5	Σ Y _{..2} = 15 Ȳ _{..2} = 5	Σ Y _{..3} = 14 Ȳ _{..3} = 4,66	Σ Y _{...} = 44 Ȳ _{...} = 4,88	

Es necesario advertir que, en el diseño que nos ocupa, la cantidad de tratamientos, de sujetos y de órdenes o períodos es la misma. En las fórmulas incluidas en la Tabla 9.26 se ha utilizado el término *a* para indicar dicha cantidad.

TABLA 9.26 Fórmulas necesarias para el cálculo de las sumas cuadráticas del ANOVA en un diseño de cuadrado latino intrasujeto

Efecto debido a los sujetos	$SC_{\text{subj.}} = \left[\frac{1}{a} \sum_i \left(\sum_j \sum_k Y_{ijk} \right)^2 \right] - C$	(9.122)
Efecto debido al orden o período	$SC_{\text{per.}} = \left[\frac{1}{a} \sum_k \left(\sum_i \sum_j Y_{ijk} \right)^2 \right] - C$	(9.123)
Efecto debido a los tratamientos	$SC_{\text{trat.}} = \left[\frac{1}{a} \sum_j \left(\sum_i \sum_k Y_{ijk} \right)^2 \right] - C$	(9.124)
Variabilidad total	$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - C$	(9.125)
Variabilidad residual o del error	$SC_{\text{res.}} = SCT - SC_{\text{subj.}} - SC_{\text{per.}} - SC_{\text{trat.}}$	(9.126)
C	$C = \frac{1}{a^2} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2$	(9.127)

Tras obtener el valor de C, llevaremos a cabo el cálculo de las diferentes sumas de cuadrados.

$$C = \frac{1}{3^2} (44)^2 = 215,11$$

Una vez obtenido este valor, procedemos a calcular la $SC_{\text{suj.}}$, la $SC_{\text{per.}}$, la $SC_{\text{trat.}}$, la SCT y la $SC_{\text{res.}}$, respectivamente:

$$SC_{\text{suj.}} = \frac{1}{3} [(16)^2 + (15)^2 + (13)^2] - 215,11 = 216,66 - 215,11 = 1,55$$

$$SC_{\text{per.}} = \frac{1}{3} [(15)^2 + (15)^2 + (14)^2] - 215,11 = 215,33 - 215,11 = 0,22$$

$$SC_{\text{trat.}} = \frac{1}{3} [(19)^2 + (14)^2 + (11)^2] - 215,11 = 226 - 215,11 = 10,89$$

$$SCT = [(7)^2 + (5)^2 + (3)^2 + (5)^2 + (4)^2 + (6)^2 + (4)^2 + (6)^2 + (4)^2] - 215,11 = \\ = 228 - 215,11 = 12,89$$

$$SC_{\text{res.}} = 12,89 - 1,55 - 0,22 - 10,89 = 0,23$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 9.27 Análisis de la varianza para el diseño intrasujeto de cuadrado latino 3×3 : ejemplo práctico

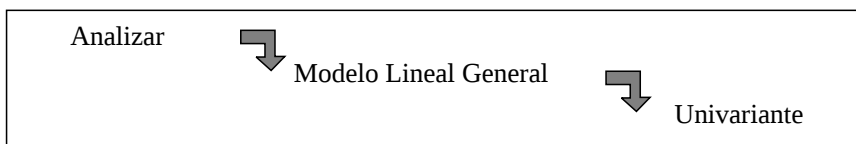
Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Sujetos	1,55	$a - 1 = 2$	0,77	$F_{\text{suj.}} = 7$
Períodos	0,22	$a - 1 = 2$	0,11	$F_{\text{per.}} = 1$
Tratamientos	10,89	$a - 1 = 2$	5,44	$F_{\text{trat.}} = 49,45$
Residual	0,23	$(a - 1)(a - 2) = 2$	0,11	
Total	12,89	$a^2 - 1 = 8$		

Tras consultar las *tablas de los valores críticos de la distribución F* con un nivel de confianza del 95 % ($\alpha = 0,05$) y trabajando con una hipótesis de una cola, podemos concluir que el *nivel de familiaridad de las palabras (factor A)* ejerce una influencia estadísticamente significativa sobre su posterior *recuerdo* ($F_{0,95; 2,2} = 19 < F_{\text{obs.}} = 49,45$). Por otra parte, mantenemos la hipótesis nula tanto para el *efecto debido a los sujetos* ($F_{0,95; 2,2} = 19 > F_{\text{obs.}} = 7$) como para el *derivado del orden en el que se administran los tratamientos* ($F_{0,95; 2,2} = 19 > F_{\text{obs.}} = 1$).

• Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la(s) variable(s) dependiente(s) y los factores, tal y como se ha descrito en el Epígrafe 8.2.2.4, referido al análisis de datos mediante el paquete estadístico SPSS 10.0 en los diseños de cuadrado latino.



- El menú **Modelo** nos permite especificar el modelo mediante el que se va a realizar el análisis que, por defecto, es un modelo factorial completo. En nuestro caso, debemos escoger la opción *Personalizado*. A continuación, nos debemos situar en *Construir términos* y seleccionar la opción *Efectos princip.* del menú desplegable. Tras escoger los tres factores de nuestro ejemplo, debemos colocarlos en el cuadro derecho donde se indica *Modelo*, utilizando el botón destinado a tal fin:



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar, de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar «estimaciones del tamaño del efecto»* y *«potencia observada»*).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

UNIANOVA

```

recuerdo BY fam sujeto orden
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(fam)
/EMMEANS = TABLES(sujeto)
/EMMEANS = TABLES(orden)
/PRINT = ETASQ OPOWER
/CRITERIA = ALPHA(.05)
/DESIGN = fam sujeto orden.

```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
Nivel de familiaridad	1,00	Bajo	3
	2,00	Medio	3
	3,00	Alto	3
SUJETO	1,00		3
	2,00		3
	3,00		3
ORDEN	1,00	Orden 1	3
	2,00	Orden 2	3
	3,00	Orden 3	3

Pruebas de los efectos intersujetos

Variable dependiente: Recuerdo

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parám. de no centralidad	Potencia observada ^a
Modelo corregido	12,667 ^b	6	2,111	19,000	0,051	0,983	114,000	0,639
Intersección	215,111	1	215,111	1.936,0	0,001	0,999	1.936,000	1,000
FAM	10,889	2	5,444	49,000	0,020	0,980	98,000	0,918
SUJETO	1,556	2	0,778	7,000	0,125	0,875	14,000	0,331
ORDEN	0,222	2	0,111	1,000	0,500	0,500	2,000	0,096
Error	0,222	2	0,111					
Total	228,000	9						
Total corregida	12,889	8						

^a Calculado con alfa = 0,05.

^b R cuadrado = 0,983 (R cuadrado corregida = 0,931).

Medias marginales estimadas

1. Nivel de familiaridad

Variable dependiente: Recuerdo

Nivel de familiaridad	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Bajo	6,333	0,192	5,505	7,161
Medio	4,667	0,192	3,839	5,495
Alto	3,667	0,192	2,839	4,495

2. Sujeto

Variable dependiente: Recuerdo

Sujeto	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
1,00	5,333	0,192	4,505	6,161
2,00	5,000	0,192	4,172	5,828
3,00	4,333	0,192	3,505	5,161

3. Orden

Variable dependiente: Recuerdo

Orden	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Orden 1	5,000	0,192	4,172	5,828
Orden 2	5,000	0,192	4,172	5,828
Orden 3	4,667	0,192	3,839	5,495

OTRAS MODALIDADES DE DISEÑO

En el último capítulo del presente texto, abordaremos tres tipos de diseños que se caracterizan por la incorporación u omisión de determinados elementos en su estructura básica. Tales diseños son los *diseños multivariados*, los *diseños de bloques incompletos* y los *diseños fraccionados*. Como ya se ha visto en el Capítulo 4, tanto el diseño de bloques incompletos como el diseño fraccionado son diseños estructuralmente incompletos. Ambos tipos de diseños se utilizan con el objetivo de reducir el tamaño de aquellos experimentos factoriales en los que los niveles de los factores y las dimensiones de variación del diseño hacen que tales experimentos resulten difíciles de llevar a cabo.

10.1. DISEÑO EXPERIMENTAL MULTIVARIADO

10.1.1. Características generales del diseño experimental multivariado

El *diseño experimental multivariado* es una estructura de investigación en la que se registran dos o más variables dependientes de naturaleza cualitativamente distinta que, en conjunto, representan una entidad teórica compleja a la que denominamos constructo.

Como señalan Arnau (1990b) y Spector (1993), los fenómenos sociales no son tan simples como para reducirlos a características directamente observables sin mediación de una teoría. Por ello, la investigación psicológica debe recurrir a elaboraciones teóricas o constructos teóricos que contemplen la complejidad del comportamiento humano. En este sentido, desde una perspectiva metodológica, limitar la observación de una conducta a una sola variable dependiente puede acarrear problemas de validez. Es en este contexto, como bien señala López (1995), donde el diseño multivariado adquiere una relevancia esencial, ya que permite examinar los efectos de determinados tratamientos sobre comportamientos complejos, elaborados teóricamente y operacionalizados a través de un conjunto de variables dependientes.

En esencia, los diseños experimentales multivariados son una extensión de los diseños univariados. De hecho, la única diferencia que existe entre ambos modelos, en el proceso de estimación y ajuste de los parámetros, radica en el tipo de criterio estadístico utilizado para la prueba de la hipótesis que, en el caso del diseño multivariado, es una función de una matriz de sumas de cuadrados asociadas y no asociadas al modelo propuesto (véase el Epígrafe 5.2 referido al análisis multivariado de la varianza).

Desde una perspectiva metodológica, gran cantidad de autores (Arnau, 1990a; Huberty y Morris, 1989; Stevens, 1992; Tabachnick y Fidell, 1989; Vallejo y Fernández, 1992; Vallejo y Herrero, 1991) consideran que los diseños multivariados proporcionan importantes **ventajas** frente a los diseños univariados. Entre tales ventajas, cabe destacar las siguientes:

- Permiten llevar a cabo investigaciones más acordes con la complejidad del comportamiento humano.
- Permiten establecer relaciones complejas entre diferentes medidas o variables dependientes y, por tanto, proporcionan más información que las estructuras univariadas.
- Son diseños que poseen mayor sensibilidad o potencia probatoria y mayor validez que los diseños univariados.

10.1.2. El análisis de datos en el diseño multivariado: análisis multivariado de la varianza (MANOVA)

10.1.2.1. Consideraciones generales acerca del modelo multivariante

A diferencia de los diseños intersujetos, los *diseños univariados intrasujeto* pueden ser tratados desde una perspectiva analítica multivariada (véase el Epígrafe 9.1.2). No obstante, el diseño univariado implica una sola medida que, aunque pueda replicarse varias veces, no constituye un constructo y, por tanto, no es equiparable al diseño multivariado. Según Stevens (1992), el *diseño multivariado de medidas repetidas* puede definirse como una estructura de investigación en la que se toman, de cada uno de los sujetos, una serie de medidas de variables cualitativamente distintas en diferentes períodos temporales o bajo la acción de varios tratamientos y condiciones diferentes. Al igual que el diseño intrasujeto univariado, el diseño multivariado de medidas repetidas también puede ser abordado desde una perspectiva analítica univariante o multivariante. En el primer caso, nos encontramos con el modelo que se conoce como *modelo mixto multivariado de medidas repetidas*. Si, por el contrario, tanto la naturaleza conceptual del diseño como la técnica utilizada para el análisis de datos son multivariados, el modelo se denomina *modelo doblemente multivariado* (López y Ato, 1994b). Dado que este último modelo se enmarca dentro de la estrategia longitudinal aplicada, consideramos que su análisis excede los objetivos del presente texto. En consecuencia, aquí se abordará el MANOVA a partir de los datos obtenidos mediante un diseño experimental intersujetos de naturaleza transversal. Cabe señalar que nuestra finalidad básica radica en proporcionar al lector un ejemplo que le permita obtener una visión general de la forma en la que se lleva a cabo la prueba de la hipótesis en un diseño experimental multivariado, por cuya razón se ha escogido un formato de diseño sencillo para desarrollar dicho análisis. No obstante, el lector interesado en profundizar en el desarrollo del análisis multivariado de la varianza en diseños intrasujeto de múltiples variables dependientes registradas a lo largo de múltiples ocasiones o puntos temporales, puede acceder a él consultando el texto de Arnau (1995e). Por otra parte, cabe recordar que la lógica y los principales elementos de la prueba de significación multivariante ya han sido descritos en el Epígrafe 5.2 del presente texto.

10.1.2.2. Ejemplo práctico

Supongamos que en el ámbito de la psicología sexual, realizamos una investigación con el objetivo de examinar la influencia que ejerce la *edad* (factor A) sobre el *nivel de utilización de los métodos anticonceptivos* (variable dependiente 1) y sobre la *satisfacción en las relaciones sexuales* (variable dependiente 2). A tal fin, se seleccionan aleatoriamente tres sujetos pertenecientes a cada uno de los siguientes rangos de edad: (a_1) *entre 20 y 30 años*, (a_2) *entre 31 y 40 años* y (a_3) *entre 41 y 50 años*. A los nueve sujetos de la muestra se les aplica, en primer lugar, un cuestionario sobre *hábitos de uso de diferentes métodos anticonceptivos*. A continuación, todos ellos responden a una escala tipo Likert dirigida a examinar su *nivel de satisfacción en las relaciones sexuales*. En la Tabla 10.1 pueden observarse los resultados obtenidos en el estudio.

TABLA 10.1 Matriz de datos del experimento

			A (Edad de los sujetos)			
	a_1 (Entre 20 y 30 años)		a_2 (Entre 31 y 40 años)		a_3 (Entre 41 y 50 años)	
	Y_1 (Uso de métodos anticonceptivos)	Y_2 (Satisfacción sexual)	Y_1 (Uso de métodos anticonceptivos)	Y_2 (Satisfacción sexual)	Y_1 (Uso de métodos anticonceptivos)	Y_2 (Satisfacción sexual)
	8	8	11	5	6	3
	6	7	7	6	5	2
	7	6	9	4	4	4
Sumatorios y medias marginales	$\Sigma Y_{.11} = 21$ $\bar{Y}_{.11} = 7$	$\Sigma Y_{.12} = 21$ $\bar{Y}_{.12} = 7$	$\Sigma Y_{.21} = 27$ $\bar{Y}_{.21} = 9$	$\Sigma Y_{.22} = 15$ $\bar{Y}_{.22} = 5$	$\Sigma Y_{.31} = 15$ $\bar{Y}_{.31} = 5$	$\Sigma Y_{.32} = 9$ $\bar{Y}_{.32} = 3$

La Tabla 10.1 refleja la estructura correspondiente a un diseño multigrupos aleatorios multivariante. Tomando como referencia la matriz de datos de dicha tabla, procedemos a desarrollar el análisis multivariado de la varianza.

Los modelos matemáticos que subyacen al análisis estadístico bajo el supuesto de la hipótesis alternativa en el diseño que nos ocupa, son similares en el caso del ANOVA y del MANOVA. La única diferencia entre ambos radica en que mientras que en el ANOVA cada componente del modelo se representa en una sola columna, en el MANOVA disponemos de tantas columnas como variables dependientes para cada uno de los componentes de la ecuación estructural. En otras palabras, en el ANOVA partimos de vectores, mientras que en el MANOVA las sumas de cuadrados se calculan a partir de matrices. De esta forma, el modelo estructural del diseño multigrupos aleatorios multivariante puede formularse mediante la siguiente expresión:

$$Y_{ijp} = M_p + A_{jp} + E_{ijp} \quad (10.1)$$

donde:

Y_{ijp} = Puntuación obtenida en la variable dependiente p por el sujeto i bajo el j -ésimo nivel de la variable de tratamiento.

M_p = Media común a todas las observaciones en la variable dependiente p .

A_{jp} = Efecto que ejerce el j -ésimo nivel de la variable de tratamiento sobre la variable dependiente p .

E_{ijp} = Componente de error específico asociado al sujeto i , al j -ésimo nivel de la variable de tratamiento y a la variable dependiente p .

En el ejemplo que nos ocupa, las *matrices asociadas a los parámetros de la ecuación estructural* expresada mediante la Fórmula 10.1, adoptan los siguientes valores:

$$[Y] = [M] + [A] + [E]$$

$$\begin{bmatrix} 8 & 8 \\ 6 & 7 \\ 7 & 6 \\ 11 & 5 \\ 7 & 6 \\ 9 & 4 \\ 6 & 3 \\ 5 & 2 \\ 4 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \\ 7 & 5 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \\ 2 & 0 \\ 2 & 0 \\ 2 & 0 \\ -2 & -2 \\ -2 & -2 \\ -2 & -2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ 2 & 0 \\ -2 & 1 \\ 0 & -1 \\ 1 & 0 \\ 0 & -1 \\ -1 & 1 \end{bmatrix}$$

Una vez obtenidas estas matrices, podemos calcular las *matrices de sumas de cuadrados y de productos cruzados asociadas a las fuentes de variación intersujetos e intrasujeto del diseño*, multiplicando cada matriz por su traspuesta, a saber:

$$\begin{aligned} [A]^T[A] &= \begin{bmatrix} 0 & 0 & 0 & 2 & 2 & 2 & -2 & -2 & -2 \\ 2 & 2 & 2 & 0 & 0 & 0 & -2 & -2 & -2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & 2 \\ 0 & 2 \\ 2 & 0 \\ 2 & 0 \\ 2 & 0 \\ -2 & -2 \\ -2 & -2 \\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 24 & 12 \\ 12 & 24 \end{bmatrix} = \\ &= \begin{bmatrix} SC_{A(Y_1)} & SP_{A(Y_1Y_2)} \\ SP_{A(Y_1Y_2)} & SC_{A(Y_2)} \end{bmatrix} \end{aligned} \quad (10.2)$$

$$\begin{aligned} [E]^T[E] &= \begin{bmatrix} 1 & -1 & 0 & 2 & -2 & 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 & 1 & -1 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ 2 & 0 \\ -2 & 1 \\ 0 & -1 \\ 1 & 0 \\ 0 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 12 & -2 \\ -2 & 6 \end{bmatrix} = \\ &= \begin{bmatrix} SC_{e(Y_1)} & SP_{e(Y_1Y_2)} \\ SP_{e(Y_1Y_2)} & SC_{e(Y_2)} \end{bmatrix} \end{aligned} \quad (10.3)$$

Como se ha señalado en el Apartado 5.2 referido al análisis multivariado de la varianza, la *lambda de Wilks* (Λ) ha sido y sigue siendo el estadístico más utilizado para llevar a cabo la prueba de la hipótesis en el MANOVA (Bray y Maxwell, 1993). Por tal razón, en el presente ejemplo calcularemos el *valor de Λ para la fuente de variación asociada al efecto de tratamiento*:

$$\Lambda_A = \frac{|SC_{\text{error}}|}{|SC_{\text{total}}|} = \frac{\begin{vmatrix} 12 & -2 \\ -2 & 6 \end{vmatrix}}{\left[\begin{vmatrix} 24 & 12 \\ 12 & 24 \end{vmatrix} + \begin{vmatrix} 12 & -2 \\ -2 & 6 \end{vmatrix} \right]} = \frac{\begin{vmatrix} 12 & -2 \\ -2 & 6 \end{vmatrix}}{\begin{vmatrix} 36 & 10 \\ 10 & 30 \end{vmatrix}} = \frac{68}{980} = 0,069 \quad (10.4)$$

Una vez hallado el valor de Λ , se puede estimar el valor de una *variable cuya distribución se aproxima a la distribución F con v_1 y v_2 grados de libertad*, aplicando la Fórmula 10.5.

$$\hat{F} = \frac{1 - \Lambda^{1/t}}{\Lambda^{1/t}} \frac{v_2}{v_1} \quad (10.5)$$

donde:

$v_1 = p \cdot gl_{\text{intersujetos}}$, siendo p = número de variables dependientes.

$v_2 = gl_{\text{intersujetos}} (gl_{\text{error}} - p + 1)$.

$$t = \begin{cases} 1 & \text{si } v_1 = 2 \\ \sqrt{\frac{v_1^2 - 4}{p^2 + gl_{\text{intersujetos}}^2 - 5}} & \text{si } v_1 \neq 2 \end{cases}$$

Aplicando dicha fórmula a los datos de nuestro ejemplo, obtenemos:

$$v_1 = 2 \cdot 2 = 4$$

$$v_2 = 2(6 - 2 + 1) = 10$$

$$t = \sqrt{\frac{4^2 - 4}{2^2 + 2^2 - 5}} = \sqrt{\frac{12}{3}} = 2$$

Por tanto:

$$\hat{F}_A = \frac{1 - (0,069)^{1/2}}{(0,069)^{1/2}} \cdot \frac{10}{4} = \frac{0,737}{0,263} = 2,799$$

Llegados a este punto, podemos elaborar la tabla del análisis multivariado de la varianza.

TABLA 10.2 Análisis multivariado de la varianza para un diseño multigrupos aleatorios multivariante: ejemplo práctico

Estadístico	Valor	V_1	V_2	Razón $\hat{F}$	p
Lambda de Wilks	0,069	4	10	7,005	< 0,05

Dado que el valor observado o empírico del estadístico ($\hat{F}_{\text{observada}} = 7,005$) es mayor que su valor crítico o teórico ($F_{\text{crit.}(0,05; 4,10)} = 3,48$), rechazamos el modelo asociado con la hipótesis de nulidad. En consecuencia, cabe concluir que existen diferencias estadísticamente significativas en el uso de los métodos anticonceptivos (Y_1) y en la satisfacción en las relaciones sexuales (Y_2), en función de la edad (factor A).

• **Estimación de la lambda de Wilks a partir de los autovalores asociados a la matriz R_F**

Como ya se ha señalado en el epígrafe referido al análisis multivariado de la varianza (véase el punto 5.2), la estimación de la razón F en dicho análisis requiere calcular la matriz análoga al cociente $SC_{\text{intersujetos}}/SC_{\text{error}}$ del ANOVA o la matriz que se conoce como matriz R_F . Esta matriz se obtiene multiplicando la inversa de la matriz W , o matriz correspondiente a la fuente de variación residual, por la matriz B o matriz correspondiente al efecto del tratamiento.

$$R_F = W^{-1}B = \frac{B}{W} \quad (10.6)$$

En el ejemplo que nos ocupa:

$$R_F = \begin{bmatrix} 12 & -2 \\ -2 & 6 \end{bmatrix}^{-1} \begin{bmatrix} 24 & 12 \\ 12 & 24 \end{bmatrix}$$

Calculemos la matriz inversa de la matriz W :

$$\begin{aligned} W^{-1} &= \begin{bmatrix} 12 & -2 \\ -2 & 6 \end{bmatrix}^{-1} = \frac{1}{\begin{vmatrix} 12 & -2 \\ -2 & 6 \end{vmatrix}} \begin{bmatrix} 6(-1)^{1+1} & -2(-1)^{1+2} \\ -2(-1)^{2+1} & 12(-1)^{2+2} \end{bmatrix} = \\ &= \frac{1}{68} \begin{bmatrix} 6 & 2 \\ 2 & 12 \end{bmatrix} = \begin{bmatrix} \frac{3}{34} & \frac{1}{34} \\ \frac{1}{34} & \frac{6}{34} \end{bmatrix} \end{aligned}$$

Por tanto:

$$R_F = \begin{bmatrix} \frac{3}{34} & \frac{1}{34} \\ \frac{1}{34} & \frac{6}{34} \end{bmatrix} \begin{bmatrix} 24 & 12 \\ 12 & 24 \end{bmatrix} = \begin{bmatrix} \frac{42}{17} & \frac{30}{17} \\ \frac{48}{17} & \frac{78}{17} \end{bmatrix}$$

Para interpretar esta matriz se calculan los *autovalores* o «*eigenvalues*» asociados a ella, resolviendo la denominada *ecuación característica*. Partiendo de que al sustraer el autovalor a los elementos de la diagonal principal (sumas de cuadrados de cada una de las

variables dependientes) de la matriz R_F el determinante es nulo, cabe plantear la siguiente igualdad:

$$|R_F - \lambda I| = 0 \quad (10.7)$$

En el ejemplo que nos ocupa:

$$\begin{aligned}
 |R_F - \lambda I| &= 0 \\
 \left| \begin{bmatrix} \frac{42}{17} & \frac{30}{17} \\ \frac{48}{17} & \frac{78}{17} \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| &= 0 \\
 \left| \begin{bmatrix} \frac{42}{17} & \frac{30}{17} \\ \frac{48}{17} & \frac{78}{17} \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| &= 0 \\
 \left| \begin{bmatrix} \left(\frac{42}{17}\right) - \lambda & \frac{30}{17} \\ \frac{48}{17} & \left(\frac{78}{17}\right) - \lambda \end{bmatrix} \right| &= 0 \\
 \left(\frac{42}{17} - \lambda\right)\left(\frac{78}{17} - \lambda\right) - \left(\frac{48}{17} \cdot \frac{30}{17}\right) &= 0 \\
 \left(\frac{42}{17} - \lambda\right)\left(\frac{78}{17} - \lambda\right) - \frac{1.440}{289} &= 0 \\
 \frac{3.276}{289} - \frac{78\lambda}{17} - \frac{42\lambda}{17} + \lambda^2 - \frac{1.440}{289} &= 0 \\
 \lambda^2 - \frac{120\lambda}{17} + \frac{108}{17} &= 0 \\
 17\lambda^2 - 120\lambda + 108 &= 0
 \end{aligned}$$

Recordemos que la fórmula para la resolución de la ecuación de segundo grado es:

$$\begin{aligned}
 ax^2 + bx + c &= 0 \\
 x &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (10.8)
 \end{aligned}$$

Aplicando la Fórmula (10.8) a los datos de nuestro ejemplo, obtenemos λ_1 y λ_2 o los autovalores asociados a la matriz R_F .

$$\lambda = \frac{-(-120) \pm \sqrt{(-120)^2 - 4(17)(108)}}{2(17)} = \frac{120 \pm \sqrt{7.056}}{34}$$

$$\lambda_1 = \frac{120 + 84}{34} = 6$$

$$\lambda_2 = \frac{120 - 84}{34} = \frac{18}{17}$$

A continuación, debemos comprobar que tanto λ_1 como λ_2 cumplen con la ecuación característica.

$$\begin{vmatrix} \frac{42}{17} - 6 & \frac{30}{17} \\ \frac{48}{17} & \frac{78}{17} - 6 \end{vmatrix} = 0 \Rightarrow \left(\frac{42 - 102}{17} \right) \left(\frac{78 - 102}{17} \right) - \left(\frac{30}{17} \cdot \frac{48}{17} \right) = \frac{1.440}{17^2} - \frac{1.440}{17^2} = 0$$

$$\begin{vmatrix} \frac{42}{17} - \frac{18}{17} & \frac{30}{17} \\ \frac{48}{17} & \frac{78}{17} - \frac{18}{17} \end{vmatrix} = 0 \Rightarrow \left(\frac{24}{17} \cdot \frac{60}{17} \right) - \left(\frac{30}{17} \cdot \frac{48}{17} \right) = \frac{1.440}{17^2} - \frac{1.440}{17^2} = 0$$

El número de autovalores que cumplen con la ecuación característica se simboliza mediante la letra s , y es igual al menor de los valores correspondientes a la cantidad de grados de libertad intersujetos y a la cantidad de variables dependientes. En nuestro caso:

$$s = \min. (gl_{\text{intersujetos}}, p) = \min. (a - 1, p) = \min. (2, 2) = 2$$

La propiedad más relevante de los autovalores es que nos permiten obtener el valor de la *lambda de Wilks*. Dicho índice o estadístico se calcula mediante la siguiente expresión:

$$\Lambda = \pi_{i=1}^s \frac{1}{1 + \lambda_i} \quad (10.9)$$

Aplicando la Fórmula (10.9) a los datos de nuestro ejemplo, obtenemos:

$$\Lambda = \frac{1}{1 + 6} \cdot \frac{1}{1 + \frac{18}{17}} = \frac{1}{7} \cdot \frac{1}{\frac{35}{17}} = \frac{1}{7} \cdot \frac{17}{35} = \frac{17}{245} = 0,069$$

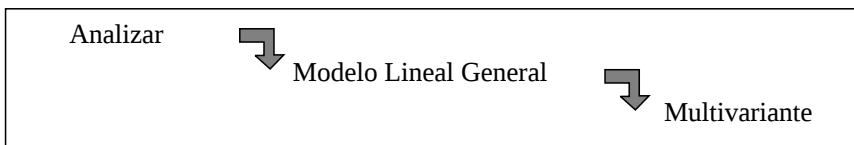
Una vez hallado el valor de Λ , no nos resta sino estimar el valor de la razón $\hat{F}_A$ y elaborar la tabla del MANOVA a fin de comprobar si la variable de tratamiento ejerce una influencia estadísticamente significativa sobre las variables dependientes. Dado que tales cálculos ya han sido realizados anteriormente, no volveremos a llevarlos a cabo.

Antes de finalizar con el presente epígrafe, cabe señalar que los autovalores asociados a la matriz R_F también permiten calcular el resto de los índices o estadísticos que se utilizan para contrastar la hipótesis de nulidad multivariante, es decir, la *traza de Pillai-Bartlett*, la *traza de Hotelling-Lawley* y la *raíz mayor de Roy*. Al igual que la *lambda de Wilks*, cualquiera de estos índices permite obtener una variable cuya distribución se aproxima a la distribución F y comparar el valor observado del estadístico con su distribución muestral bajo la hipótesis nula. El lector interesado en acceder a dichas aproximaciones puede encontrarlas en el texto de Pascual, García y Frías (1995). Por otra parte, en ese mismo libro se ofrecen diversos ejemplos en los que la prueba de significación multivariante se desarrolla mediante el denominado análisis discriminante, el cual constituye un procedimiento de amplio uso en el ámbito de las ciencias del comportamiento. Por tal razón, remitimos al lector interesado en profundizar en el MANOVA a dicho texto.

10.1.2.3. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Multivariante** del análisis **Modelo Lineal General**.



- Indicamos las variables dependientes y los factores.



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar*

«estadísticos descriptivos», «estimaciones del tamaño del efecto» y «potencia observada»).

- La sintaxis del análisis de la varianza que corresponde a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

```
GLM
  métodos satisfac BY edad
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(edad)
/PRINT = DESCRIPTIVE ETASQ OPOWER
/CRITERIA = ALPHA(.05)
/DESIGN = edad .
```

- Resultados:

Modelo lineal general

Factores intersujetos

		Etiqueta del valor	N
EDAD	1,00	20-30 años	3
	2,00	31-40 años	3
	3,00	41-50 años	3

Estadísticos descriptivos

	EDAD	Media	Desv. típ.	N
METODOS Uso de métodos anticonceptivos	1,00 20-30 años	7,0000	1,0000	3
	2,00 31-40 años	9,0000	2,0000	3
	3,00 41-50 años	5,0000	1,0000	3
	Total	7,0000	2,1213	9
SATISFAC Satisfacción sexual	1,00 20-30 años	7,0000	1,0000	3
	2,00 31-40 años	5,0000	1,0000	3
	3,00 41-50 años	3,0000	1,0000	3
	Total	5,0000	1,9365	9

Contrastes multivariados^d

Efecto		Valor	F	GI de la hipótesis	GI del error	Sig.	Eta ²	Parám. de no central.	Potencia observada ^a
Intercept	Traza de Pillai	0,990	242,87 ^b	2,000	5,000	0,000	0,990	485,735	1,000
	Lambda de Wilks	0,010	242,87 ^b	2,000	5,000	0,000	0,990	485,735	1,000
	Traza de Hotelling	97,147	242,87 ^b	2,000	5,000	0,000	0,990	485,735	1,000
	Raíz mayor de Roy	97,147	242,87 ^b	2,000	5,000	0,000	0,990	485,735	1,000

Contrastes multivariados^d (continuación)

Efecto		Valor	F	GI de la hipótesis	GI del error	Sig.	Eta ²	Parám. de no central.	Potencia observada ^a
EDAD	Traza de Pillai	1,371	6,545	4,000	12,00	0,005	0,686	26,182	0,937
	Lambda de Wilks	0,069	6,991 ^b	4,000	10,00	0,006	0,737	27,963	0,934
	Traza de Hotelling	7,059	7,059	4,000	8,000	0,010	0,779	28,235	0,903
	Raíz mayor de Roy	6,000	18,000 ^c	2,000	6,000	0,003	0,857	36,000	0,985

^a Calculado con alfa = 0,05.^b Estadístico exacto.^c El estadístico es un límite superior para la F el cual ofrece un límite inferior para el nivel de significación.^d Diseño: Intercept + EDAD.**Pruebas de los efectos intersujetos**

Fuente	Variable dependiente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta ²	Parám. de no central.	Potencia observada ^a
Modelo corregido	Uso de métodos anticonceptivos Satisfacción sexual	24,000 ^b	2	12,000	6,000	0,037	0,667	12,000	0,660
		24,000 ^c	2	12,000	12,00	0,008	0,800	24,000	0,922
Intercept	Uso de métodos anticonceptivos Satisfacción sexual	441,000	1	441,000	220,5	0,000	0,974	220,500	1,000
		225,000	1	225,000	225,0	0,000	0,974	225,000	1,000
EDAD	Uso de métodos anticonceptivos Satisfacción sexual	24,000	2	12,000	6,000	0,037	0,667	12,000	0,660
		24,000	2	12,000	12,00	0,008	0,800	24,000	0,922
Error	Uso de métodos anticonceptivos Satisfacción sexual	12,000	6	2,000					
		6,000	6	1,000					
Total	Uso de métodos anticonceptivos Satisfacción sexual	477,000	9						
		255,000	9						

Pruebas de los efectos intersujetos (*continuación*)

Fuente	Variable dependiente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta ²	Parám. de no central.	Potencia observada ^a
Total corregida	Uso de métodos anticonceptivos	36,000	8						
	Satisfacción sexual	30,000	8						

^a Calculado con alfa = 0,05.
^b R cuadrado = 0,667 (R cuadrado corregida = 0,556).
^c R cuadrado = 0,800 (R cuadrado corregida = 0,733).

Medias marginales estimadas

EDAD

Variable dependiente	EDAD	Media	Error típ.	Intervalo de confianza al 95 %	
				Límite inferior	Límite superior
Uso de métodos anticonceptivos	20-30 años	7,000	0,816	5,002	8,998
	31-40 años	9,000	0,816	7,002	10,998
	41-50 años	5,000	0,816	3,002	6,998
Satisfacción sexual	20-30 años	7,000	0,577	5,587	8,413
	31-40 años	5,000	0,577	3,587	6,413
	41-50 años	3,000	0,577	1,587	4,413

10.2. DISEÑO DE BLOQUES INCOMPLETOS

10.2.1. Características generales del diseño de bloques incompletos

Si se trabaja con un diseño factorial que consta de varios factores, se puede generar una gran cantidad de grupos de tratamiento, de manera que la formación de bloques con un número suficiente de sujetos como para que puedan ser asignados al azar a cada uno de tales grupos puede resultar muy costosa. Así, podemos encontrarnos ante una situación en la que quizá no se disponga de suficientes sujetos como para que cada bloque constituya una replicación completa del experimento.

Por otra parte, aun cuando el experimentador contase con la cantidad necesaria de sujetos para elaborar tales bloques, la construcción de bloques de gran tamaño incrementaría la heterogeneidad intrabloque y, en consecuencia, reduciría la precisión en la estimación de los efectos experimentales.

Ante dichos inconvenientes, cabe recurrir a la *técnica de confusión*, cuyo principio básico consiste en obtener bloques de menor tamaño que no constituyen replicaciones completas

del experimento, *sacrificando* la estimación de los efectos de determinadas interacciones de escasa importancia. Los diseños asociados a la técnica de confusión se conocen como *diseños de bloques incompletos*. Dado que en tales estructuras de investigación la variabilidad entre bloques queda eliminada del error experimental, son diseños eficaces para la estimación de aquellos efectos que el experimentador considere más relevantes.

10.2.2. La técnica de confusión

La técnica de confusión es un procedimiento que consiste en aplicar un bloqueo basado en bloques de menor tamaño que aquellos que constituyen replicaciones completas del experimento. El principio básico de este procedimiento radica en sacrificar determinadas interacciones de escasa importancia, incluyendo en cada bloque únicamente las combinaciones de tratamiento que presentan un mismo signo en el contraste o en la combinación lineal correspondiente a tales interacciones.

En caso de que en las sucesivas réplicas del experimento se sacrifique la misma interacción, se dice que la *confusión es completa*. Ello implica la pérdida de toda la información relativa a dicha interacción. Si, por el contrario, en cada una de las diferentes repeticiones del experimento se sacrifica una interacción distinta, el procedimiento empleado se conoce como *técnica de confusión parcial*. En este caso, sólo se dispone de una información parcial acerca de tales interacciones. A continuación, abordamos brevemente cada una de estas dos modalidades de confusión.

10.2.2.1. Técnica de confusión completa

Supongamos que deseamos llevar a cabo una investigación aplicando un diseño factorial $2 \times 2 \times 2$ y que, debido a las dimensiones de variación del diseño, consideramos adecuado utilizar la técnica de confusión completa. Como ya es sabido, el diseño 2^3 permite estimar siete efectos factoriales (A , B , C , AB , AC , BC , ABC), cuyos contrastes o combinaciones lineales son ortogonales entre sí. En cada uno de estos contrastes, la mitad de los grupos o de las combinaciones de tratamiento presenta signo positivo, mientras que a la otra mitad le corresponde signo negativo.

Procedamos a estimar los *efectos factoriales* y las *combinaciones lineales del diseño* $2 \times 2 \times 2$ tomando como referencia las medias de las diferentes combinaciones de tratamiento.

En la Tabla 10.3 se representan dichas combinaciones de tratamiento, empleando la notación de Yates (1937).

TABLA 10.3 Representación de las ocho combinaciones de tratamiento que configuran el diseño factorial $2 \times 2 \times 2$ utilizando la notación de Yates (1937)

		C			
		c_1		c_2	
B		b_1	b_2	b_1	b_2
A	a_1	(1)	(b)	(c)	(bc)
	a_2	(a)	(ab)	(ac)	(abc)

Tomando como referencia la Tabla 10.3 estimamos, en primer lugar, los *efectos principales* de A, B, y C. Para ello, calculamos el promedio de los *efectos simples* correspondientes a cada uno de tales factores.

$$\text{Efecto principal de A} = \frac{1}{4} [(a) - (1) + (ab) - (b) + (ac) - (c) + (abc) - (bc)] \quad (10.10)$$

$$\text{Efecto principal de B} = \frac{1}{4} [(b) - (1) + (ab) - (a) + (bc) - (c) + (abc) - (ac)] \quad (10.11)$$

$$\text{Efecto principal de C} = \frac{1}{4} [(c) - (1) + (ac) - (a) + (bc) - (b) + (abc) - (ab)] \quad (10.12)$$

A continuación estimamos los *efectos asociados a las interacciones de primer orden* AB, AC y BC.

Efecto de la interacción AB a través de los valores de C:

$$\text{Interacción AB} = \frac{1}{4} [(1) + (ab) - (a) - (b) + (c) + (abc) - (ac) - (bc)] \quad (10.13)$$

Efecto de la interacción AC a través de los valores de B:

$$\text{Interacción AC} = \frac{1}{4} [(1) + (ac) - (a) - (c) + (b) + (abc) - (ab) - (bc)] \quad (10.14)$$

Efecto de la interacción BC a través de los valores de A:

$$\text{Interacción BC} = \frac{1}{4} [(1) + (bc) - (c) - (b) + (a) + (abc) - (ac) - (ab)] \quad (10.15)$$

Por último, estimamos el *efecto asociado a la interacción de segundo orden*, ABC. Para ello, partimos de las dos estimaciones de cualquiera de las interacciones de primer orden respecto a cada uno de los niveles del factor que no forma parte de la interacción, y hallamos la diferencia entre ambas estimaciones. Por ejemplo, la diferencia entre la estimación de la interacción AB para c_2 y la estimación de la interacción AB para c_1 permite obtener el efecto factorial ABC.

$$\begin{aligned} \text{Interacción ABC} &= \frac{1}{4} [(c) + (abc) - (ac) - (bc)] - [(1) + (ab) - (a) - (b)] = \\ &= \frac{1}{4} [(c) + (abc) - (ac) - (bc) - (1) - (ab) + (a) + (b)] \end{aligned} \quad (10.16)$$

En la Tabla 10.4 se representan los elementos necesarios para el cálculo de las combinaciones lineales y de los efectos factoriales del diseño factorial $2 \times 2 \times 2$.

TABLA 10.4 Elementos necesarios para el cálculo de las combinaciones lineales y de los efectos factoriales del diseño factorial $2 \times 2 \times 2$

Efecto	Coeficientes para las medias de combinaciones de tratamientos								Combinación lineal	Estimación del efecto factorial promedio
	(1)	(a)	(b)	(ab)	(c)	(ac)	(bc)	(abc)		
Prin. A	—	+	—	+	—	+	—	+	C_A	$C_A/4$
Prin. B	—	—	+	+	—	—	+	+	C_B	$C_B/4$
Prin. C	—	—	—	—	+	+	+	+	C_C	$C_C/4$
Int. AB	+	—	—	+	+	—	—	+	C_{AB}	$C_{AB}/4$
Int. AC	+	—	+	—	—	+	—	+	C_{AC}	$C_{AC}/4$
Int. BC	+	+	—	—	—	—	+	+	C_{BC}	$C_{BC}/4$
Int. ABC	—	+	+	—	+	—	—	+	C_{ABC}	$C_{ABC}/4$

Tras hallar las combinaciones o los contrastes lineales y los efectos factoriales que configuran el diseño $2 \times 2 \times 2$, ya podemos aplicar la técnica de confusión completa a dicho diseño. Dado que, como señala Finney (1963), a excepción de aquellas situaciones en las que se dispone de una gran cantidad de bloques, la información que se obtiene a partir de la interacción de segundo orden $A \times B \times C$ es escasa, se suele sacrificar la estimación del efecto factorial asociado a ella a fin de trabajar con bloques de menor tamaño que los que se utilizan al aplicar la técnica de bloqueo. Para ello, se parte de la combinación lineal correspondiente a la interacción $A \times B \times C$ (véase la Tabla 10.4):

$$(a) + (b) + (c) + (abc) - (1) - (ab) - (ac) - (bc)$$

A partir de esta combinación o contraste lineal se forman dos bloques, asignando al primero de ellos las combinaciones de tratamiento con signo positivo (a, b, c, abc) y al segundo las de signo negativo ($1, ab, ac, bc$). Como es obvio, en tal disposición experimental las diferencias entre ambos bloques coinciden con el contraste lineal asociado a la interacción de segundo orden $A \times B \times C$. Suponiendo que el experimento se replica cuatro veces y que en cada una de tales réplicas se asignan cuatro sujetos a cada bloque, el plan experimental se representaría de la siguiente forma (véase la Tabla 10.5).

TABLA 10.5 Disposición experimental correspondiente a un diseño de bloques incompletos $2 \times 2 \times 2$ con la interacción $A \times B \times C$ totalmente confundida

		Replicaciones			
		1	2	3	4
Bloques	1	a b c abc	a b c abc	a b c abc	a b c abc
	2	1 ab ac bc	1 ab ac bc	1 ab ac bc	1 ab ac bc

Cabe señalar que mediante esta disposición experimental se elimina del error experimental la variabilidad entre bloques de cuatro sujetos experimentales. No obstante, si se hubiese aplicado un diseño de bloques totalmente aleatorios, se hubiese eliminado la variabilidad correspondiente a ocho unidades experimentales.

En la Tabla 10.6 se representa la descomposición de los grados de libertad del análisis de la varianza para un diseño de bloques incompletos $2 \times 2 \times 2$ con la interacción $A \times B \times C$ totalmente confundida:

TABLA 10.6 Descomposición de los grados de libertad del ANOVA para un diseño de bloques incompletos $2 \times 2 \times 2$ con la interacción $A \times B \times C$ totalmente confundida

Fuentes de variación	Grados de libertad
Efecto principal de A	$a - 1$
Efecto principal de B	$b - 1$
Efecto principal de C	$c - 1$
Efecto de interacción $A \times B$	$(a - 1)(b - 1)$
Efecto de interacción $A \times C$	$(a - 1)(c - 1)$
Efecto de interacción $B \times C$	$(b - 1)(c - 1)$
Bloques	$2r - 1$
Error (residual)	$6r - 6$
Total	$8r - 1$

Donde r es la cantidad de réplicas del experimento.

10.2.2.2. *Técnica de confusión parcial*

En el subapartado anterior (véase la sección correspondiente a la *técnica de confusión completa*) hemos observado que en la estimación del efecto factorial asociado a la interacción $A \times B \times C$ todos los componentes que presentaban signo positivo se hallaban en bloques distintos a los que presentaban signo negativo, es decir, que el efecto de dicha interacción estaba totalmente confundido con la variación entre bloques a lo largo de las sucesivas réplicas del diseño. Sin embargo, en las diferentes repeticiones de un experimento, también es posible confundir, alternativamente, diferentes efectos factoriales. En eso consiste, de hecho, la estrategia que se conoce como *técnica de confusión parcial*. Así, siguiendo con el formato de diseño que hemos utilizado para ilustrar la técnica de confusión completa, si realizáramos cuatro replicaciones de un diseño factorial $2 \times 2 \times 2$, asignando cuatro sujetos experimentales a cada bloque, el experimento podría disponerse de tal forma que, en cada replicación, la variabilidad entre bloques quedara confundida con el efecto factorial asociado a una interacción diferente, a saber, el efecto de la interacción $A \times B \times C$ con la variabilidad entre bloques de la primera replicación, el de la interacción $A \times B$ con la de la segunda, el de la interacción $A \times C$ con la de la tercera y el de la interacción $B \times C$ con la variación entre bloques de la cuarta replicación. En tal caso, el plan experimental se representaría de la siguiente manera (véase la Tabla 10.7):

TABLA 10.7 Disposición experimental correspondiente a un diseño de bloques incompletos $2 \times 2 \times 2$ con las interacciones $A \times B \times C$, $A \times B$, $A \times C$ y $B \times C$ parcialmente confundidas

		Replicaciones			
		1	2	3	4
Bloques	1	a b c abc	a b ac bc	a c ab bc	b c ab ac
	2	1 ab ac bc	1 c ab abc	1 b ac abc	1 a bc abc
Efecto de confusión		$A \times B \times C$	$A \times B$	$A \times C$	$B \times C$

Como señala Arnau (1986), cuando se aplica la técnica de confusión parcial a los datos del experimento, el *análisis estadístico* resulta algo más complicado que cuando se emplea la técnica de confusión completa. Así, en el primer caso, debemos considerar por separado cada una de las $(2^n - 1)$ combinaciones lineales y efectos factoriales que configuran el diseño¹. En primer lugar, se calculan las sumas cuadráticas de los efectos factoriales que no están confundidos (en el ejemplo que nos ocupa, los efectos principales de los factores A , B y C). Cada uno de los restantes efectos de confusión ha de ser estimado a partir de aquellas replicaciones en las que dicho efecto no ha sido confundido. Así, por ejemplo, la estimación intrabloque de la suma de cuadrados de la interacción $A \times B \times C$ se obtiene a partir de las réplicas 2, 3 y 4 (véase la Tabla 10.7). De esta forma, dado que cada uno de estos efectos puede estimarse a partir de aquellas replicaciones cuya diferencia entre los totales de los bloques no coincide con la combinación lineal asociada a dicho efecto, cabe afirmar que las cuatro interacciones del diseño están parcialmente confundidas con diferencias específicas entre bloques. En definitiva, para cada uno de tales efectos, realizamos una estimación intrabloque basada en tres de las cuatro réplicas del experimento.

En lo que respecta al ANOVA del diseño de bloques incompletos con interacciones parcialmente confundidas, cabe señalar que como cada una de las interacciones se confunde con una determinada replicación del experimento, sólo se dispone de información parcial para cada una de ellas. Esta información, que en nuestro ejemplo se representa mediante la razón 3/4, recibe el nombre de *información relativa de las interacciones confundidas* y se define como la relación que existe entre la cantidad de información que se posee de un efecto parcialmente confundido y de un efecto no confundido.

Como cabe deducir de lo que se ha afirmado a lo largo de la presente sección, las sumas de cuadrados de los efectos no confundidos se estiman a partir de todas las repeticiones del diseño. Sin embargo, la variabilidad asociada a cada uno de los efectos confundidos se obtiene a partir de las replicaciones en las que dicho efecto no ha sido confundido. En

¹ El lector interesado en el proceso de estimación de tales contrastes y efectos en el caso del diseño $2 \times 2 \times 2$, puede consultarlos en la sección referida a la *técnica de confusión completa*.

consecuencia, las estimaciones de estos últimos efectos son menos precisas que las de los primeros.

En la Tabla 10.8 se representa la descomposición de los grados de libertad del análisis de la varianza para un diseño de bloques incompletos $2 \times 2 \times 2$ con las interacciones $A \times B \times C$, $A \times B$, $A \times C$ y $B \times C$ parcialmente confundidas.

TABLA 10.8 Descomposición de los grados de libertad del ANOVA para un diseño de bloques incompletos $2 \times 2 \times 2$ con las interacciones $A \times B \times C$, $A \times B$, $A \times C$ y $B \times C$ parcialmente confundidas

Fuentes de variación	Grados de libertad
Efecto principal de A	$a - 1$
Efecto principal de B	$b - 1$
Efecto principal de C	$c - 1$
Efecto de interacción $A \times B$	$(a - 1)(b - 1)$
Efecto de interacción $A \times C$	$(a - 1)(c - 1)$
Efecto de interacción $B \times C$	$(b - 1)(c - 1)$
Efecto de interacción $A \times B \times C$	$(a - 1)(b - 1)(c - 1)$
Bloques	$2r - 1$
Error (residual)	$6r - 7$
Total	$8r - 1$

Donde r es la cantidad de réplicas del experimento.

Tras describir a grandes rasgos en qué consisten las técnicas de confusión completa y parcial, desarrollaremos el ANOVA en un diseño de bloques incompletos de formato sencillo. Dado que nuestra finalidad básica consiste en proporcionar una perspectiva general de esta modalidad de diseño, aplicaremos un solo procedimiento para llevar a cabo el análisis. El lector interesado en profundizar en estas disposiciones experimentales puede consultar, entre otros, los textos clásicos de Cochran y Cox (1957) y de Finney (1963), así como los trabajos de Bose y Nair (1939) y de Yates (1936). Por otra parte, el libro de Arnau (1986) ofrece diversas estrategias para calcular las combinaciones lineales y los efectos factoriales de distintas modalidades de diseño, lo que puede resultar útil para aplicar las técnicas de confusión completa y parcial que caracterizan a los diseños de bloques incompletos.

10.2.3. El análisis de la varianza para el diseño de bloques incompletos

10.2.3.1. Ejemplo práctico

Supongamos que en el ámbito de la psicología clínica, realizamos una investigación con el objetivo de examinar la influencia que ejerce el *sexo* (factor A) y la *ubicación del barrio en el que se reside* (factor B) sobre la *fobia a la oscuridad* (variable dependiente). No obstante,

se considera que una posible variable extraña capaz de confundir los resultados del estudio es *el hecho de haber sido, o no, víctima de una agresión sexual*. Con el objeto de controlar dicha variable, y teniendo presente lo difícil que resulta conseguir sujetos experimentales que cumplan con el criterio de bloqueo, se opta por utilizar un diseño de bloques incompletos, sacrificando la información asociada a la interacción $A \times B$, la cual se considera teóricamente irrelevante. De esta forma, se dividen los sujetos en dos bloques de 2 sujetos cada uno, en función de *haber sido* (bloque 1) o *no haber sido* (bloque 2) *víctima de una agresión sexual*. Por otra parte, los sujetos del primer bloque son *mujeres* (a_1) *residentes en barrios periféricos* (b_1), ($a_1b_1 = 1$), u *hombres* (a_2) *residentes en barrios del centro urbano* (b_2), ($a_2b_2 = ab$), mientras que los sujetos del segundo bloque son *hombres* (a_2) *residentes en barrios periféricos* (b_1), ($a_2b_1 = a$), o *mujeres* (a_1) *residentes en barrios del centro urbano* (b_2), ($a_1b_2 = b$). En la Tabla 10.9 pueden observarse los resultados obtenidos por los sujetos en una prueba destinada a medir la fobia a la oscuridad.

TABLA 10.9 Matriz de datos del experimento

		Replicaciones			Totales por bloque
		1	2	3	
Bloques (Víctima de agresión sexual)	1 (Sí)	5($a_1b_1 = 1$) 4($a_2b_2 = ab$)	4($a_1b_1 = 1$) 5($a_2b_2 = ab$)	6($a_1b_1 = 1$) 4($a_2b_2 = ab$)	15 13
	2 (No)	2($a_2b_1 = a$) 7($a_1b_2 = b$)	4($a_2b_1 = a$) 5($a_1b_2 = b$)	3($a_2b_1 = a$) 6($a_1b_2 = b$)	9 18
Totales por replicación		18	18	19	55

En el diseño que nos ocupa, las combinaciones lineales asociadas al efecto principal de A , al efecto principal de B y a la interacción $A \times B$ se definen mediante las siguientes expresiones:

$$\text{Combinación lineal asociada al efecto principal de } A = [- (1) + (a) - (b) + (ab)] \quad (10.17)$$

$$\text{Combinación lineal asociada al efecto principal de } B = [- (1) - (a) + (b) + (ab)] \quad (10.18)$$

$$\text{Combinación lineal asociada al efecto principal de } A \times B = [+ (1) - (a) - (b) + (ab)] \quad (10.19)$$

A fin de aplicar la estrategia de confusión total, la combinación lineal asociada a la interacción $A \times B$ ha de coincidir con la diferencia entre ambos bloques. Para ello, se deben colocar las combinaciones de tratamiento con signo positivo ($1, ab$) en el primer bloque y las de signo negativo (a, b) en el segundo. De esta forma, el efecto de la interacción $A \times B$ se halla totalmente confundido con la variación o diferencia entre bloques a lo largo de las sucesivas réplicas del diseño (véase la disposición experimental en la Tabla 10.9).

Dicho esto, procedemos a desarrollar el análisis de la varianza para el diseño de bloques incompletos 2×2 con la interacción $A \times B$ totalmente confundida, tomando como referencia la matriz de datos de la Tabla 10.9.

10.2.3.2. Desarrollo del análisis de la varianza para el diseño de bloques incompletos 2×2 con la interacción $A \times B$ totalmente confundida

• Variabilidad total

$$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{1}{N} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2 \quad (10.20)$$

Donde:

Y_{ijk} = Puntuación obtenida en la variable dependiente por el i -ésimo sujeto en el j -ésimo bloque bajo la k -ésima replicación del experimento.

N = Cantidad total de observaciones.

Por tanto:

$$SCT = [(5)^2 + (4)^2 + (2)^2 + (7)^2 + (4)^2 + (5)^2 + (4)^2 + (5)^2 + (6)^2 + (4)^2 + (3)^2 + (6)^2] - \frac{(55)^2}{12}$$

$$SCT = 273 - 252,08 = 20,92$$

• Efecto del factor A

$$SCA = \frac{C_A^2}{N} = \frac{1}{12} (-15 + 9 - 18 + 13)^2 = \frac{1}{12} (-11)^2 = 10,08 \quad (10.21)$$

Donde:

C_A = Combinación lineal asociada al efecto principal del factor A. Dicho componente se obtiene multiplicando cada coeficiente correspondiente a una determinada combinación de tratamientos por la puntuación total obtenida en la misma y sumando entre sí tales productos. En el caso del factor A: $(-1)(15) + (1)(9) + (-1)(18) + (1)(13) = -11$ (véase la Fórmula (10.17)).

• Efecto del factor B

$$SCB = \frac{C_B^2}{N} = \frac{1}{12} (-15 - 9 + 18 + 13)^2 = \frac{1}{12} (7)^2 = 4,08 \quad (10.22)$$

Donde el componente adicional C_B hace referencia a la combinación lineal asociada al efecto principal del factor B (véase la Fórmula (10.18)).

• Variabilidad asociada a los bloques

$$SC_{\text{Bloques}} = \left[\sum_j \sum_k \left(\sum_i Y_{ijk} \right)^2 / n \right] - \frac{1}{N} \left(\sum_i \sum_j \sum_k Y_{ijk} \right)^2 \quad (10.23)$$

Donde:

n = Cantidad de observaciones por bloque en cada réplica.

Por tanto:

$$SC_{\text{Bloques}} = \left[\frac{9^2}{2} + \frac{9^2}{2} + \frac{9^2}{2} + \frac{9^2}{2} + \frac{10^2}{2} + \frac{9^2}{2} \right] - \frac{(55)^2}{12} = 252,5 - 252,08 = 0,42$$

• **Variabilidad residual o del error**

$$SCE = SCT - SCA - SCB - SC_{\text{Bloques}} \quad (10.24)$$

Por tanto:

$$SCE = 20,92 - 10,08 - 4,08 - 0,42 = 6,34$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 10.10 Análisis de la varianza para el diseño de bloques incompletos 2×2 con la interacción $A \times B$ totalmente confundida: ejemplo práctico

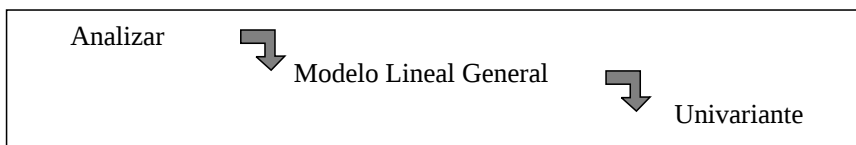
Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Efecto principal de A	10,08	1	10,08	6,38
Efecto principal de B	4,08	1	4,08	2,58
Bloques	0,42	5	0,08	0,05
Error (residual)	6,34	4	1,58	
Total	20,92	11		

Tras consultar las *tablas de los valores críticos de la distribución F* con un nivel de confianza del 95 % ($\alpha = 0,05$) y trabajando con una hipótesis de una cola, podemos concluir que ni el sexo (factor A) ($F_{0,95; 1,4} = 7,71 > F_{\text{obs.}} = 6,38$) ni la *ubicación del barrio en el que se reside* (factor B) ($F_{0,95; 1,4} = 7,71 > F_{\text{obs.}} = 2,58$) ejercen una influencia estadísticamente significativa sobre la *fobia que presentan los sujetos a la oscuridad*. De la misma forma, el *hecho de haber sido, o no, víctima de una agresión sexual* (factor de bloqueo) tampoco afecta significativamente a la variable criterio ($F_{0,95; 5,4} = 6,26 > F_{\text{obs.}} = 0,05$).

10.2.3.3. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor SPSS 10.0.

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la variable dependiente (en nuestro ejemplo, la variable *fobia a la oscuridad*) y los factores (en nuestro ejemplo, los factores *sexo*, *ubicación del barrio de residencia*, *bloques* y *número de replicación*).
- A continuación, mediante el menú **Modelo**, debemos especificar el modelo con el que se va a realizar el análisis ya que, por defecto, el programa utiliza un modelo factorial completo. En nuestro caso, hemos de escoger la opción *Personalizado* y construir los términos tal y como se ha descrito en el Epígrafe 8.2.2.4. En el ejemplo que nos ocupa, las fuentes de variación que debemos especificar son los efectos principales de los dos factores de nuestro diseño (*sexo* y *ubicación del barrio de residencia*) así como la interacción entre el factor de bloqueo (víctima de agresión sexual) y la variable replicaciones.



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar, de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar* «estadísticos descriptivos», «estimaciones del tamaño del efecto» y «potencia observada»).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

UNIANOVA

```

fobia BY sexo barrio bloques replica
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(sexo)
/EMMEANS = TABLES(barrio)

```

```

/EMMEANS = TABLES(bloques*replica)
/PRINT = DESCRIPTIVE ETASQ OPOWER
/CRITERIA = ALPHA(.05)
/DESIGN = sexo barrio bloques*replica .

```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		Etiqueta del valor	N
SEXO	1	Mujeres	6
	2	Varones	6
Ubicación del barrio de residencia	1	Periférico	6
	2	Centro	6
Víctima de agresión sexual	1	Sí	6
	2	No	6
Replicaciones	1		4
	2		4
	3		4

Estadísticos descriptivos

Variable dependiente: Fobia a la oscuridad

SEXO	Ubicación del barrio de resid.	Víctima de agresión sexual	Replicaciones	Media	Desv. tip.	N
Mujeres	Periférico	Sí	1	5,00	,	1
			2	4,00	,	1
			3	6,00	,	1
			Total	5,00	1,00	3
		Total	1	5,00	,	1
			2	4,00	,	1
			3	6,00	,	1
			Total	5,00	1,00	3
	Centro	No	1	7,00	,	1
			2	5,00	,	1
			3	6,00	,	1
			Total	6,00	1,00	3
		Total	1	7,00	,	1
			2	5,00	,	1
			3	6,00	,	1
			Total	6,00	1,00	3
	Total	Sí	1	5,00	,	1
			2	4,00	,	1
			3	6,00	,	1
			Total	5,00	1,00	3
		No	1	7,00	,	1
			2	5,00	,	1
			3	6,00	,	1
			Total	6,00	1,00	3
		Total	1	6,00	1,41	2
			2	4,50	0,71	2
			3	6,00	0,00	2
			Total	5,50	1,05	6

Estadísticos descriptivos (continuación)

Variable dependiente: Fobia a la oscuridad

SEXO	Ubicación del barrio de resid.	Víctima de agresión sexual	Replicaciones	Media	Desv. tip.	N
Varones	Periférico	No	1	2,00	,	1
			2	4,00	,	1
			3	3,00	,	1
			Total	3,00	1,00	3
		Total	1	2,00	,	1
			2	4,00	,	1
			3	3,00	,	1
			Total	3,00	1,00	3
	Centro	Sí	1	4,00	,	1
			2	5,00	,	1
			3	4,00	,	1
			Total	4,33	0,58	3
		Total	1	4,00	,	1
			2	5,00	,	1
			3	4,00	,	1
			Total	4,33	0,58	3
	Total	Sí	1	4,00	,	1
			2	5,00	,	1
			3	4,00	,	1
			Total	4,33	0,58	3
		No	1	2,00	,	1
			2	4,00	,	1
			3	3,00	,	1
			Total	3,00	1,00	3
		Total	1	3,00	1,41	2
			2	4,50	0,71	2
			3	3,50	0,71	2
			Total	3,67	1,03	6
Total	Periférico	Sí	1	5,00	,	1
			2	4,00	,	1
			3	6,00	,	1
			Total	5,00	1,00	3
		No	1	2,00	,	1
			2	4,00	,	1
			3	3,00	,	1
			Total	3,00	1,00	3
		Total	1	3,50	2,12	2
			2	4,00	0,00	2
			3	4,50	2,12	2
			Total	4,00	1,41	6
	Centro	Sí	1	4,00	,	1
			2	5,00	,	1
			3	4,00	,	1
			Total	4,33	0,58	3
		No	1	7,00	,	1
			2	5,00	,	1
			3	6,00	,	1
			Total	6,00	1,00	3
		Total	1	5,50	2,12	2
			2	5,00	0,00	2
			3	5,00	1,41	2
			Total	5,17	1,17	6
	Total	Sí	1	4,50	0,71	2
			2	4,50	0,71	2
			3	5,00	1,41	2
			Total	4,67	0,82	6
		No	1	4,50	3,54	2
			2	4,50	0,71	2
			3	4,50	2,12	2
			Total	4,50	1,87	6
		Total	1	4,50	2,08	4
			2	4,50	0,58	4
			3	4,75	1,50	4
			Total	4,58	1,38	12

Pruebas de los efectos intersujetos**Variable dependiente: Fobia a la oscuridad**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no central.	Potencia observada ^a
Modelo corregido	14,583 ^b	7	2,083	1,316	0,417	0,697	9,211	0,182
Intersección	252,083	1	252,083	159,21	0,000	0,975	159,211	1,000
SEXO	10,083	1	10,083	6,368	0,065	0,614	6,368	0,484
BARRIO	4,083	1	4,083	2,579	0,184	0,392	2,579	0,237
BLOQUES * REPLICA	0,417	5	8,333E-02	0,053	0,997	0,062	0,263	0,054
Error	6,333	4	1,583					
Total	273,000	12						
Total corregida	20,917	11						

^a Calculado con alfa = 0,05.^b R cuadrado = 0,697 (R cuadrado corregida = 0,167).**Medias marginales estimadas****1. Sexo****Variable dependiente: Fobia a la oscuridad**

Sexo	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Mujeres	5,500	0,514	4,074	6,926
Varones	3,667	0,514	2,240	5,093

2. Ubicación del barrio de residencia**Variable dependiente: Fobia a la oscuridad**

Ubicación del barrio de residencia	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
Periférico	4,000	0,514	2,574	5,426
Centro	5,167	0,514	3,740	6,593

3. Víctima de agresión sexual * Replicaciones

Variable dependiente: Fobia a la oscuridad

Víctima de agresión sexual	Replicaciones	Media	Error típ.	Intervalo de confianza al 95 %	
				Límite inferior	Límite superior
Sí	1	4,500	0,890	2,030	6,970
	2	4,500	0,890	2,030	6,970
	3	5,000	0,890	2,530	7,470
No	1	4,500	0,890	2,030	6,970
	2	4,500	0,890	2,030	6,970
	3	4,500	0,890	2,030	6,970

10.3. DISEÑO FACTORIAL FRACCIONADO

10.3.1. Características generales del diseño factorial fraccionado

En la medida en que las dimensiones de variación de un experimento aumentan, también se incrementa la cantidad de grupos de tratamiento necesaria para llevarlo a cabo. Así, puede ocurrir que el tamaño del experimento exceda las posibilidades de realizarlo empleando un diseño factorial completo. Al igual que los diseños de bloques incompletos, que ya han sido abordados en el Epígrafe 10.2, los diseños factoriales fraccionados proporcionan una solución a dicho problema. Por esta razón, en la actualidad son ampliamente utilizados en el ámbito de la psicología de las organizaciones y, en especial, en el análisis del control de calidad del personal, el cual depende de una gran cantidad de factores que deben ser depurados a fin de esclarecer cuáles son los que realmente ejercen una influencia relevante sobre dicho control.

Los *diseños factoriales fraccionados*, conocidos también como *diseños de replicación fraccionada*, son estructuras de investigación en las que sólo se utiliza una parte de la totalidad de las combinaciones de tratamientos en cada una de las réplicas del experimento. Por tanto, incluyen una *fracción* de todas las posibles combinaciones de tratamientos que se derivarían del uso de un diseño factorial completo. Finney (1963) afirma que, mientras la técnica de confusión se caracteriza por la reducción del tamaño del bloque, la replicación fraccionada extiende dicho principio a todo el experimento. En este sentido, el diseño factorial fraccionado consigue reducir la cantidad de grupos experimentales pero, en contrapartida, algunas de las interacciones del experimento no pueden ser estimadas y, además, se produce una confusión entre determinados efectos factoriales. En consecuencia, esta clase de diseño debe partir del supuesto de que las interacciones entre los tratamientos son nulas ya que, en caso contrario, la interpretación de los resultados sería incorrecta. López (1995) añade que el posible confundido entre las fuentes de variación especificadas y no especificadas en el diseño puede afectar considerablemente a la validez de la investigación.

Con respecto a los *orígenes del diseño factorial fraccionado*, cabe señalar que esta estructura de investigación fue inicialmente propuesta por Finney (1945, 1946) para diseños factoriales simples del tipo 2^k y 3^k . No obstante, el mérito de sistematizar y de generalizar la utilización de la técnica de replicación fraccionada con diseños más complejos del tipo p^k , siendo p cualquier número primo y k la cantidad de factores, corresponde a Kempthorne (1947, 1952).

Kirk (1982) proporciona un excelente resumen de las *situaciones en las que*, a su juicio, *el uso de los diseños factoriales fraccionados resulta especialmente apropiado*. En su opinión, tales diseños deben emplearse en los siguientes casos:

- En aquellos experimentos en los que se incluye una gran cantidad de factores.
- En aquellos diseños en los que todas las variables de tratamiento poseen la misma cantidad de niveles.
- En aquellas investigaciones en las que existen razones teóricas para poder asumir como nulas las interacciones de orden superior.
- En investigaciones exploratorias y en situaciones en las que se requiere la evaluación y el seguimiento de la efectividad de una serie de experimentos secuencialmente encadenados. En tales situaciones, en las que se suele comenzar realizando un experimento en el que se incluyen gran cantidad de factores, el diseño factorial fraccionado proporciona la posibilidad de planificar nuevos experimentos con los factores que hayan resultado más efectivos en el experimento inicial.

Arnau (1996) destaca que las *fracciones que se utilizan con mayor frecuencia* suelen ser descritas como potencias de $1/2$ y suelen aplicarse a diseños factoriales a dos niveles. Por tanto, los diseños que se fraccionan habitualmente son diseños factoriales de base 2 o diseños 2^k , de cuyo fraccionamiento se obtienen diseños factoriales de replicación fraccionada de una mitad ($1/2$), de un cuarto [$(1/2)^2 = 1/4$], de un octavo [$(1/2)^3 = 1/8$], etc. Sin embargo, la técnica de replicación fraccionada puede también aplicarse a diseños factoriales con una misma cantidad de niveles y cuya base sea un número primo (3, 5, 7, etc.). Así, con diseños factoriales de base tres, a saber, 3^k , es posible derivar réplicas fraccionadas de un tercio ($1/3$), de un noveno [$(1/3)^2 = 1/9$], de un veintisieteavo [$(1/3)^3 = 1/27$], etc. En el supuesto de que la cantidad de niveles no sea igual a un número primo, el experimentador puede transformar el diseño en una estructura factorial del tipo 2^2 , 2^3 , 3^2 , etc., y, posteriormente, fraccionarlo.

Para finalizar con las características generales de estas estructuras de investigación, cabe señalar que Box, Hunter y Hunter (1978) han realizado una excelente sistematización acerca de las características más relevantes de los diseños factoriales fraccionados 2^{k-p} , donde k hace referencia a la cantidad de factores y p a la potencia a la que se eleva la fracción $1/2$. Aunque no vamos a describir esta clase de diseños, cabe apuntar que suelen clasificarse en función del grado de fraccionamiento que se aplica en ellos y que, habitualmente, suelen categorizarse en tres modalidades, a saber, Diseños de resolución III, Diseños de resolución IV y Diseños de resolución V.

Tras haber descrito las características generales de los diseños factoriales fraccionados, abordaremos brevemente la técnica de replicación fraccionada y, a continuación, desarrollaremos el análisis de la varianza en un diseño factorial fraccionado de formato sencillo. Dado que nuestra finalidad básica consiste en proporcionar una perspectiva general de esta modalidad de diseño, aplicaremos un solo procedimiento para llevar a cabo el análisis. El lector interesado en profundizar en el conocimiento de estas disposiciones experimentales puede consultar, entre otros, los textos de Anderson y McLean (1974) y de Kempthorne (1952) y los trabajos de Finney (1945, 1946, 1963).

10.3.2. La técnica de replicación fraccionada

Ilustraremos la *técnica de replicación fraccionada* tomando como ejemplo un diseño factorial a dos niveles (2^k), ya que los principios que se derivan de esta modalidad de di-

seño son fácilmente generalizables a cualquier otra estructura factorial. La cantidad de combinaciones de tratamientos necesaria para llevar a cabo una replicación de una mitad de un diseño factorial a dos niveles se obtiene mediante la expresión $(1/2) \cdot 2^k$. Así, por ejemplo, si el diseño consta de $k = 3$ factores, la replicación fraccionada de una mitad viene dada por $(1/2) \cdot 2^3 = 2^{-1} \cdot 2^3 = 2^2 = 4$, es decir, implica utilizar la mitad (cuatro) de las combinaciones que generaría la estructura factorial completa (ocho).

En primer lugar, debemos elaborar la *matriz de los coeficientes de las combinaciones lineales del diseño* 2^3 (véase la Tabla 10.11).

TABLA 10.11 Matriz de los coeficientes de las combinaciones lineales del diseño factorial 2^3

Grupos o combinaciones de tratamientos			Coeficientes para las medias de combinaciones de tratamientos							
Factor A	Factor B	Factor C	1	A	B	C	AB	AC	BC	ABC
a_1	b_1	$c_1 = (1)$	+	+	+	+	+	+	+	+
a_1	b_1	$c_2 = (c)$	+	+	+	-	+	-	-	-
a_1	b_2	$c_1 = (b)$	+	+	-	+	-	+	-	-
a_1	b_2	$c_2 = (bc)$	+	+	-	-	-	-	+	+
a_2	b_1	$c_1 = (a)$	+	-	+	+	-	-	+	-
a_2	b_1	$c_2 = (ac)$	+	-	+	-	-	+	-	+
a_2	b_2	$c_1 = (ab)$	+	-	-	+	+	-	-	+
a_2	b_2	$c_2 = (abc)$	+	-	-	-	+	+	+	-

Como ya es sabido, los coeficientes definen el contraste o la forma en la que se combinan los grupos para la estimación de un determinado efecto factorial. Así, por ejemplo, el efecto principal de A se estima a partir de la siguiente combinación lineal:

$$\text{Efecto principal de A} = \frac{1}{4} [(1) + (c) + (b) + (bc) - (a) - (ac) - (ab) - (abc)]$$

o bien invirtiendo los signos (véase la Fórmula (10.10) en el Epígrafe 10.2.2.1 referido a los *diseños de bloques incompletos*).

El procedimiento para el cálculo de los restantes efectos factoriales sigue la misma lógica.

Tras elaborar esta matriz, debemos escoger el efecto que va a ser «sacrificado» a fin de aplicar la técnica de replicación fraccionada. Como ya se ha señalado con anterioridad, la información que normalmente se obtiene a partir de la interacción de segundo orden $A \times B \times C$ suele ser escasa. En consecuencia, este efecto se suele utilizar para llevar a cabo el fraccionamiento del diseño en dos mitades, es decir, se utiliza como *contraste de definición*. Partiendo de tal contraste, cabe reagrupar los grupos en dos mitades en base a la igualdad de los signos. Es decir, podemos dividir el diseño factorial completo 2^3 en dos mitades o en dos subdiseños. De esta forma, se obtienen las dos replicaciones de una mitad de dicho diseño (véase la Tabla 10.12).

Como cabe apreciar en la Tabla 10.12, en la estimación de cualquier efecto, la mitad de las combinaciones de cada una de las mitades toma el signo positivo, mientras que la otra mitad adopta el signo negativo, a excepción del efecto asociado a la interacción triple, en el que todas las combinaciones del tratamiento ubicadas en cada una de las mitades del diseño presentan el mismo signo. Dado que el diseño factorial completo ha sido dividido en dos replicas fraccionadas de una mitad tomando como referencia los signos positivos y negativos de la combinación lineal asociada a la interacción $A \times B \times C$, este efecto no puede ser estimado a partir de cada una de tales replicas. En consecuencia, se sacrifica la estimación de dicho efecto factorial.

TABLA 10.12 Replicas de una mitad del diseño factorial 2^3

Grupos o combinaciones de tratamientos			Replicación de una mitad	Coeficientes para las medias de combinaciones de tratamientos							
Factor A	Factor B	Factor C		1	A	B	C	AB	AC	BC	ABC
a_1	b_1	$c_1 = (1)$	1	+	+	+	+	+	+	+	+
a_1	b_2	$c_2 = (bc)$		+	+	-	-	-	-	+	+
a_2	b_1	$c_2 = (ac)$		+	-	+	-	-	+	-	+
a_2	b_2	$c_1 = (ab)$		+	-	-	+	+	-	-	+
a_1	b_1	$c_2 = (c)$	2	+	+	+	-	+	-	+	-
a_1	b_2	$c_1 = (b)$		+	+	-	+	-	+	+	-
a_2	b_1	$c_1 = (a)$		+	-	+	+	-	-	-	-
a_2	b_2	$c_2 = (abc)$		+	-	-	-	+	+	-	-

Por otra parte, si consideramos la primera replicación fraccionada del diseño, en la que los coeficientes del contraste lineal asociado a la interacción $A \times B \times C$ adoptan signo positivo, cabe comprobar que el patrón de los coeficientes para estimar el efecto principal de A coincide con el patrón de los coeficientes para calcular el efecto de interacción $B \times C$. Lo mismo ocurre con el efecto principal de B y la interacción $A \times C$ y con el efecto principal de C y la interacción $A \times B$. Por tanto, cabe afirmar que cada uno de los efectos factoriales principales se confunde con una de las interacciones de primer orden. En cuanto a la segunda replicación fraccionada del diseño, podemos observar que se dan los mismos patrones pero con signos diferentes. Conviene recordar que la inversión de los signos de los coeficientes de la combinación lineal no afecta a su interpretación ni al valor de la suma de cuadrados calculado a partir de ella. En consecuencia, para la segunda mitad del diseño, los patrones también están confundidos. Finney (1945) se ha referido a este efecto con el nombre de *alias*. En palabras del autor, el *alias* de un determinado efecto es el contraste con el que dicho efecto queda confundido. En el ejemplo que nos ocupa, la interacción $B \times C$ es un alias de A , la interacción $A \times C$ un alias de B y la interacción $A \times B$ un alias de C .

En un diseño factorial 2^k el alias de un determinado efecto factorial puede obtenerse aplicando un procedimiento algebraico consistente en multiplicar el efecto con el contraste de definición y en eliminar del resultado obtenido a partir de ese producto, cualquier término

elevado a la segunda potencia. Por ejemplo, si el contraste de definición del diseño factorial 2^3 es la interacción $A \times B \times C$, el alias del efecto principal de A se obtiene a partir de la siguiente operación: $(A) \times (ABC) = A^2BC = BC$. Dado que $A \times A = I$, es posible eliminar A^2 de la expresión anterior. De la misma forma, realizando las operaciones $(B) \times (ABC) = AB^2C = AC$ y $(C) \times (ABC) = AB^2C^2 = AB$, obtenemos los alias de los efectos principales de B y C .

10.3.3. El análisis de la varianza para el diseño factorial fraccionado

10.3.3.1 Ejemplo práctico

Supongamos que se desea llevar a cabo una investigación utilizando un diseño factorial $2 \times 2 \times 2$ y que, debido a las características específicas del área en la que se trabaja, la cantidad de grupos de tratamiento necesaria para llevarlo a cabo excede las posibilidades de realizarlo utilizando un diseño factorial completo. Ante esta situación, se opta por emplear un diseño factorial 2^3 con replicación fraccionada de una mitad. A fin de sacrificar la mínima información posible, se selecciona la interacción de segundo orden $A \times B \times C$ como contraste de definición. Partiendo de tal contraste, se divide el diseño en dos mitades y se lleva a cabo el experimento con las combinaciones de tratamientos pertenecientes a cualquiera de las dos replicas fraccionadas de una mitad; por ejemplo, con las combinaciones que siguiendo la notación de Yates se representan mediante las letras latinas a, b, c y abc . En la Tabla 10.13 pueden observarse las puntuaciones hipotéticas obtenidas, a partir de una muestra de 12 sujetos, en la variable dependiente.

TABLA 10.13 Matriz de datos del experimento

	Combinación de tratamientos				
	$a_2b_1c_1 = a$	$a_1b_2c_1 = b$	$a_1b_1c_2 = c$	$a_2b_2c_2 = abc$	
	8 6 5	3 2 3	9 8 10	5 6 4	
Totales por combinación de tratamientos	19	8	27	15	69

10.3.3.2. Desarrollo del análisis de la varianza para el diseño factorial fraccionado

• Variabilidad total:

$$SCT = \sum_i \sum_j Y_{ij}^2 - \frac{1}{N} \left(\sum_i \sum_j Y_{ij} \right)^2$$

(10.25)

Donde:

Y_{ij} = Puntuación obtenida en la variable dependiente por el i -ésimo sujeto bajo la j -ésima combinación de tratamientos.

N = Cantidad total de observaciones.

Por tanto:

$$SCT = [(8)^2 + (6)^2 + (5)^2 + (3)^2 + (2)^2 + (3)^2 + (9)^2 + (8)^2 + (10)^2 + (5)^2 + (6)^2 + (4)^2] - \frac{(69)^2}{12}$$

$$SCT = 469 - 396,75 = 72,25$$

• **Variabilidad asociada a los tratamientos**

$$SC_{\text{tratamientos}} = \left[\sum_j \left(\sum_i Y_{ij} \right)^2 / n \right] - \frac{1}{N} \left(\sum_i \sum_j Y_{ij} \right)^2 \quad (10.26)$$

Donde:

n = Cantidad de observaciones por combinación de tratamientos.

Por tanto:

$$SC_{\text{tratamientos}} = \left[\frac{19^2}{3} + \frac{8^2}{3} + \frac{27^2}{3} + \frac{15^2}{3} \right] - \frac{(69)^2}{12} = 459,67 - 396,75 = 62,92$$

• **Efecto del factor A**

$$SCA = \frac{C_A^2}{N} = \frac{1}{12} (-19 + 8 + 27 - 15)^2 = 0,08 \quad (10.27)$$

Donde:

C_A = Combinación lineal asociada al efecto principal del factor A. Dicho componente se obtiene multiplicando cada coeficiente correspondiente a una determinada combinación de tratamientos por la puntuación total obtenida en ella y sumando entre sí tales productos. En el caso del factor A: $(-1)(19) + (1)(8) + (1)(27) + (-1)(15) = 1$ (véase la Tabla 10.12).

• **Efecto del factor B**

$$SCB = \frac{C_B^2}{N} = \frac{1}{12} (19 - 8 + 27 - 15)^2 = 44,08 \quad (10.28)$$

Donde el componente adicional C_B hace referencia a la combinación lineal asociada al efecto principal del factor B (véase la Tabla 10.12).

• **Efecto del factor C**

$$SCC = \frac{C_C^2}{N} = \frac{1}{12} (19 + 8 - 27 - 15)^2 = 18,75 \quad (10.29)$$

Donde el componente adicional C_c hace referencia a la combinación lineal asociada al efecto principal del factor C (véase la Tabla 10.12).

• Variabilidad residual o del error

$$SCE = SCT - SC_{\text{tratamientos}} \tag{10.30}$$

Por tanto:

$$SCE = 72,25 - 62,92 = 9,33$$

Llegados a este punto, podemos elaborar la tabla del análisis de la varianza.

TABLA 10.14 Análisis de la varianza para el diseño factorial 2³ con replicación fraccionada de una mitad: ejemplo práctico

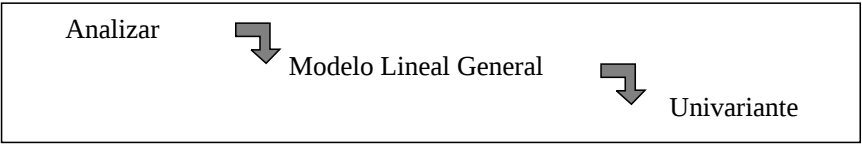
Fuentes de variación	Sumas de cuadrados	Grados de libertad	Medias cuadráticas	F
Factor A (y BC)	0,08	$a - 1 = 1$	0,08	0,07
Factor B (y AC)	44,08	$b - 1 = 1$	44,08	37,68
Factor C (y AB)	18,75	$c - 1 = 1$	18,75	16,02
Error (residual)	9,33	$ab(n - 1) = 8$	1,75	
Total	72,25	$abn - 1 = 11$		

Tras consultar las *tablas de los valores críticos de la distribución F* con un nivel de confianza del 95 % ($\alpha = 0,05$) y trabajando con una hipótesis de una cola, podemos concluir que el *factor A* no ejerce una influencia estadísticamente significativa sobre la variable dependiente ($F_{\text{obser.}} = 0,07 < F_{0,95; 1,8} = 5,32$). Sin embargo, tanto en el caso del *factor B* ($F_{\text{obser.}} = 37,68 > F_{0,95; 1,8} = 5,32$) como del *factor C* ($F_{\text{obser.}} = 16,02 > F_{0,95; 1,8} = 5,32$) se rechaza el modelo asociado a la hipótesis de nulidad, lo que permite concluir que ambos factores afectan de forma significativa a las puntuaciones obtenidas por los sujetos en la variable criterio.

10.3.3.3. Análisis de datos mediante el paquete estadístico SPSS 10.0

La forma en la que se introducen los datos puede consultarse en el Anexo B: Introducción de datos en el editor del SPSS 10.0.

- Escogemos la opción **Univariante** del análisis **Modelo Lineal General**.



- Indicamos la variable dependiente (en nuestro ejemplo VD) y los factores (en nuestro ejemplo A, B y C).
- A continuación, mediante el menú **Modelo**, debemos especificar el modelo con el que se va a realizar el análisis ya que, por defecto, el programa utiliza un modelo factorial completo. En nuestro caso, hemos de escoger la opción *Personalizado* y construir los términos tal y como se ha descrito en el Epígrafe 8.2.2.4. En el ejemplo que nos ocupa, las fuentes de variación que debemos especificar son los efectos principales de los tres factores de nuestro diseño (A, B y C).



- El menú **Opciones** proporciona la posibilidad de *mostrar las medias marginales* para cada factor. Asimismo, permite seleccionar de entre un conjunto de opciones, aquellas que deseamos que sean mostradas en los resultados (en nuestro ejemplo, se han escogido las opciones *Mostrar las medias marginales* para cada factor, así como *Mostrar* «estadísticos descriptivos», «estimaciones del tamaño del efecto» y «potencia observada»).
- La sintaxis del análisis de la varianza correspondiente a nuestro ejemplo, incluyendo las opciones arriba señaladas, sería:

UNIANOVA

```
vd BY a b c
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/EMMEANS = TABLES(a)
/EMMEANS = TABLES(b)
/EMMEANS = TABLES(c)
/PRINT = DESCRIPTIVE ETASQ OPOWER
/CRITERIA = ALPHA(.05)
/DESIGN = a b c .
```

- Resultados:

Análisis de varianza univariante

Factores intersujetos

		<i>N</i>
<i>A</i>	1,00	6
	2,00	6
<i>B</i>	1,00	6
	2,00	6
<i>C</i>	1,00	6
	2,00	6

Estadísticos descriptivos

Variable dependiente: VD

<i>A</i>	<i>B</i>	<i>C</i>	Media	Desv. típ.	<i>N</i>
1,00	1,00	2,00	9,0000	1,0000	3
		Total	9,0000	1,0000	3
	2,00	1,00	2,6667	0,5774	3
		Total	2,6667	0,5774	3
	Total	2,00	9,0000	1,0000	3
		1,00	2,6667	0,5774	3
		Total	5,8333	3,5449	6
2,00	1,00	1,00	6,3333	1,5275	3
		Total	6,3333	1,5275	3
	2,00	2,00	5,0000	1,0000	3
		Total	5,0000	1,0000	3
	Total	2,00	5,0000	1,0000	3
		1,00	6,3333	1,5275	3
		Total	5,6667	1,3663	6
Total	1,00	2,00	9,0000	1,0000	3
		1,00	6,3333	1,5275	3
		Total	7,6667	1,8619	6
	2,00	2,00	5,0000	1,0000	3
		1,00	2,6667	0,5774	3
		Total	3,8333	1,4720	6
	Total	2,00	7,0000	2,3664	6
		1,00	4,5000	2,2583	6
		Total	5,7500	2,5628	12

Pruebas de los efectos intersujetos**Variable dependiente: VD**

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Sig.	Eta cuadrado	Parámetro de no centralidad	Potencia observada ^a
Modelo corregido	62,917 ^b	3	20,972	17,976	0,001	0,871	53,929	0,999
Intersección	396,750	1	396,750	340,07	0,000	0,977	340,071	1,000
A	8,333E-02	1	8,3E-02	0,071	0,796	0,009	0,071	0,056
B	44,083	1	44,083	37,786	0,000	0,825	37,786	1,000
C	18,750	1	18,750	16,071	0,004	0,668	16,071	0,938
Error	9,333	8	1,167					
Total	469,000	12						
Total corregido	72,250	11						

^a Calculado con alfa = 0,05.^b R cuadrado = 0,871 (R cuadrado corregida = 0,822).**Medias marginales estimadas****1. A****Variable dependiente: VD**

A	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
1,00	5,833	0,441	4,816	6,850
2,00	5,667	0,441	4,650	6,684

2. B**Variable dependiente: VD**

B	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
1,00	7,667	0,441	6,650	8,684
2,00	3,833	0,441	2,816	4,850

3. C**Variable dependiente: VD**

C	Media	Error típ.	Intervalo de confianza al 95 %	
			Límite inferior	Límite superior
1,00	4,500	0,441	3,483	5,517
2,00	7,000	0,441	5,983	8,017

ANEXO



TABLAS ESTADÍSTICAS

TABLA 1 Valores críticos de la distribución t

El valor del estadístico t se encuentra en el cuerpo de la tabla, en función de los grados de libertad indicados en la columna n . La probabilidad está señalada en la fila superior de la tabla.

n	0,25	0,20	0,15	0,10	0,05	0,025	0,02	0,01	0,005	0,0025	0,001	0,0005
1	1,000	1,376	1,963	3,078	6,314	12,71	15,89	31,82	63,66	127,3	318,3	636,6
2	0,816	1,061	1,386	1,886	2,920	4,303	4,849	6,965	9,925	14,09	22,33	31,60
3	0,765	0,978	1,250	1,638	2,353	3,182	3,482	4,541	5,841	7,453	10,21	12,92
4	0,741	0,941	1,190	1,533	2,132	2,776	2,999	3,747	4,604	5,598	7,173	8,610
5	0,727	0,920	1,156	1,476	2,015	2,571	2,757	3,365	4,032	4,773	5,893	6,869
6	0,718	0,906	1,134	1,440	1,943	2,447	2,612	3,143	3,707	4,317	5,208	5,959
7	0,711	0,896	1,119	1,415	1,895	2,365	2,517	2,998	3,499	4,029	4,785	5,408
8	0,706	0,889	1,108	1,397	1,860	2,306	2,449	2,896	3,355	3,833	4,501	5,041
9	0,703	0,883	1,100	1,383	1,833	2,262	2,398	2,821	3,250	3,690	4,297	4,781
10	0,700	0,879	1,093	1,372	1,812	2,228	2,359	2,764	3,169	3,581	4,144	4,587
11	0,697	0,876	1,088	1,363	1,796	2,201	2,328	2,718	3,106	3,497	4,025	4,437
12	0,695	0,873	1,083	1,356	1,782	2,179	2,303	2,681	3,055	3,428	3,930	4,318
13	0,694	0,870	1,079	1,350	1,771	2,160	2,282	2,650	3,012	3,372	3,852	4,221
14	0,692	0,868	1,076	1,345	1,761	2,145	2,264	2,624	2,977	3,326	3,787	4,140
15	0,691	0,866	1,074	1,341	1,753	2,131	2,249	2,602	2,947	3,286	3,733	4,073
16	0,690	0,865	1,071	1,337	1,746	2,120	2,235	2,583	2,921	3,252	3,686	4,015
17	0,689	0,863	1,069	1,333	1,740	2,110	2,224	2,567	2,898	3,222	3,646	3,965
18	0,688	0,862	1,067	1,330	1,734	2,101	2,214	2,552	2,878	3,197	3,611	3,922
19	0,688	0,861	1,066	1,328	1,729	2,093	2,205	2,539	2,861	3,174	3,579	3,883
20	0,687	0,860	1,064	1,325	1,725	2,086	2,197	2,528	2,845	3,153	3,552	3,850
21	0,686	0,859	1,063	1,323	1,721	2,080	2,189	2,518	2,831	3,135	3,527	3,819
22	0,686	0,858	1,061	1,321	1,717	2,074	2,183	2,508	2,819	3,119	3,505	3,792
23	0,685	0,858	1,060	1,319	1,714	2,069	2,177	2,500	2,807	3,104	3,485	3,768
24	0,685	0,857	1,059	1,318	1,711	2,064	2,172	2,492	2,797	3,091	3,467	3,745
25	0,684	0,856	1,058	1,316	1,708	2,060	2,167	2,485	2,787	3,078	3,450	3,725
26	0,684	0,856	1,058	1,315	1,706	2,056	2,162	2,479	2,779	3,067	3,435	3,707
27	0,684	0,855	1,057	1,314	1,703	2,052	2,154	2,473	2,771	3,057	3,421	3,690
28	0,683	0,855	1,056	1,313	1,701	2,048	2,154	2,467	2,763	3,047	3,408	3,674
29	0,683	0,854	1,055	1,311	1,699	2,045	2,150	2,462	2,756	3,038	3,396	3,659
30	0,683	0,854	1,055	1,310	1,697	2,042	2,147	2,457	2,750	3,030	3,385	3,646
40	0,681	0,851	1,050	1,303	1,684	2,021	2,123	2,423	2,704	2,971	3,307	3,551
50	0,679	0,849	1,047	1,295	1,676	2,009	2,109	2,403	2,678	2,937	3,261	3,496
60	0,679	0,848	1,045	1,296	1,671	2,000	2,099	2,390	2,660	2,915	3,232	3,460
80	0,678	0,846	1,043	1,292	1,664	1,990	2,088	2,374	2,639	2,887	3,195	3,416
100	0,677	0,845	1,042	1,290	1,660	1,984	2,081	2,364	2,626	2,871	3,174	3,390
1000	0,675	0,842	1,037	1,282	1,646	1,962	2,056	2,330	2,581	2,813	3,098	3,300
inf.	0,674	0,841	1,036	1,282	1,645	1,960	2,054	2,326	2,576	2,807	3,091	3,291

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 2 Distribución F ($P = 0,001$)

Valores del estadístico F para $p = 0,001$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación ($df1$) y de la fuente de error ($df2$) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

$df2/df1$	1	2	3	4	5	6	7	8	9	10	$df1/df2$
3	167,03	148,50	141,11	137,10	134,58	132,85	131,59	130,62	129,86	129,25	3
4	74,14	61,25	56,18	53,44	51,71	50,53	49,66	49,00	48,48	48,05	4
5	47,18	37,12	33,20	31,09	29,75	28,84	28,16	27,65	27,25	26,92	5
6	35,51	27,00	23,70	21,92	20,80	20,03	19,46	19,03	18,69	18,41	6
7	29,25	21,69	18,77	17,20	16,21	15,52	15,02	14,63	14,33	14,08	7
8	25,42	18,49	15,83	14,39	13,49	12,86	12,40	12,05	11,77	11,54	8
9	22,86	16,39	13,90	12,56	11,71	11,13	10,70	10,37	10,11	9,89	9
10	21,04	14,91	12,55	11,28	10,48	9,93	9,52	9,20	8,96	8,75	10
11	19,69	13,81	11,56	10,35	9,58	9,05	8,66	8,36	8,12	7,92	11
12	18,64	12,97	10,80	9,63	8,89	8,38	8,00	7,71	7,48	7,29	12
13	17,82	12,31	10,21	9,07	8,35	7,86	7,49	7,21	6,98	6,80	13
14	17,14	11,78	9,73	8,62	7,92	7,44	7,08	6,80	6,58	6,40	14
15	16,59	11,34	9,34	8,25	7,57	7,09	6,74	6,47	6,26	6,08	15
16	16,12	10,97	9,01	7,94	7,27	6,81	6,46	6,20	5,98	5,81	16
17	15,72	10,66	8,73	7,68	7,02	6,56	6,22	5,96	5,75	5,58	17
18	15,38	10,39	8,49	7,46	6,81	6,36	6,02	5,76	5,56	5,39	18
19	15,08	10,16	8,28	7,27	6,62	6,18	5,85	5,59	5,39	5,22	19
20	14,82	9,95	8,10	7,10	6,46	6,02	5,69	5,44	5,24	5,08	20
22	14,38	9,61	7,80	6,81	6,19	5,76	5,44	5,19	4,99	4,83	22
24	14,03	9,34	7,55	6,59	5,98	5,55	5,24	4,99	4,80	4,64	24
26	13,74	9,12	7,36	6,41	5,80	5,38	5,07	4,83	4,64	4,48	26
28	13,50	8,93	7,19	6,25	5,66	5,24	4,93	4,70	4,51	4,35	28
30	13,29	8,77	7,05	6,13	5,53	5,12	4,82	4,58	4,39	4,24	30
35	12,90	8,47	6,79	5,88	5,30	4,89	4,60	4,36	4,18	4,03	35
40	12,61	8,25	6,60	5,70	5,13	4,73	4,44	4,21	4,02	3,87	40
45	12,39	8,09	6,45	5,56	5,00	4,61	4,32	4,09	3,91	3,76	45
50	12,22	7,96	6,34	5,46	4,90	4,51	4,22	4,00	3,82	3,67	50
60	11,97	7,77	6,17	5,31	4,76	4,37	4,09	3,87	3,69	3,54	60
70	11,80	7,64	6,06	5,20	4,66	4,28	3,99	3,77	3,60	3,45	70
80	11,67	7,54	5,97	5,12	4,58	4,20	3,92	3,71	3,53	3,39	80
100	11,50	7,41	5,86	5,02	4,48	4,11	3,83	3,61	3,44	3,30	100
200	11,16	7,15	5,63	4,81	4,29	3,92	3,65	3,43	3,26	3,12	200
500	10,96	7,00	5,51	4,69	4,18	3,81	3,54	3,33	3,16	3,02	500
1000	10,89	6,96	5,46	4,66	4,14	3,78	3,51	3,30	3,13	2,99	1000
>1000	1,04	6,92	5,43	4,62	4,11	3,75	3,48	3,27	3,10	2,96	>1000
$df2/df1$	1	2	3	4	5	6	7	8	9	10	$df1/df2$

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 2 Distribución F ($P = 0,001$) (Continuación)

Valores del estadístico F para $p = 0,001$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	11	12	13	14	15	16	17	18	19	df1/df2
3	128,74	128,32	127,96	127,65	127,38	127,14	126,93	126,74	126,57	3
4	47,71	47,41	47,16	46,95	46,76	46,60	46,45	46,32	46,21	4
5	26,65	26,42	26,22	26,06	25,91	25,78	25,67	25,57	25,48	5
6	18,18	17,99	17,83	17,68	17,56	17,45	17,35	17,27	17,19	6
7	13,88	13,71	13,56	13,43	13,32	13,23	13,14	13,06	12,99	7
8	11,35	11,20	11,06	10,94	10,84	10,75	10,67	10,60	10,54	8
9	9,72	9,57	9,44	9,33	9,24	9,15	9,08	9,01	8,95	9
10	8,59	8,45	8,33	8,22	8,13	8,05	7,98	7,91	7,86	10
11	7,76	7,63	7,51	7,41	7,32	7,24	7,18	7,11	7,06	11
12	7,14	7,01	6,89	6,79	6,71	6,63	6,57	6,51	6,45	12
13	6,65	6,52	6,41	6,31	6,23	6,16	6,09	6,03	5,98	13
14	6,26	6,13	6,02	5,93	5,85	5,78	5,71	5,66	5,60	14
15	5,94	5,81	5,71	5,62	5,54	5,46	5,40	5,35	5,29	15
16	5,67	5,55	5,44	5,35	5,27	5,21	5,14	5,09	5,04	16
17	5,44	5,32	5,22	5,13	5,05	4,99	4,92	4,87	4,82	17
18	5,25	5,13	5,03	4,94	4,87	4,80	4,74	4,68	4,63	18
19	5,08	4,97	4,87	4,78	4,70	4,64	4,58	4,52	4,47	19
20	4,94	4,82	4,72	4,64	4,56	4,50	4,44	4,38	4,33	20
22	4,70	4,58	4,49	4,40	4,33	4,26	4,20	4,15	4,10	22
24	4,51	4,39	4,30	4,21	4,14	4,07	4,02	3,96	3,92	24
26	4,35	4,24	4,14	4,06	3,99	3,92	3,86	3,81	3,77	26
28	4,22	4,11	4,01	3,93	3,86	3,80	3,74	3,69	3,64	28
30	4,11	4,00	3,91	3,83	3,75	3,69	3,63	3,58	3,54	30
35	3,90	3,79	3,70	3,62	3,55	3,48	3,43	3,38	3,33	35
40	3,75	3,64	3,55	3,47	3,40	3,34	3,28	3,23	3,19	40
45	3,64	3,53	3,44	3,36	3,29	3,23	3,17	3,12	3,08	45
50	3,55	3,44	3,35	3,27	3,20	3,14	3,09	3,04	2,99	50
60	3,42	3,32	3,23	3,15	3,08	3,02	2,96	2,91	2,87	60
70	3,33	3,23	3,14	3,06	2,99	2,93	2,88	2,83	2,78	70
80	3,27	3,16	3,07	3,00	2,93	2,87	2,81	2,76	2,72	80
100	3,18	3,07	2,99	2,91	2,84	2,78	2,73	2,68	2,63	100
200	3,01	2,90	2,82	2,74	2,67	2,61	2,56	2,51	2,47	200
500	2,91	2,81	2,72	2,64	2,58	2,52	2,46	2,41	2,37	500
1000	2,87	2,77	2,69	2,61	2,54	2,48	2,43	2,38	2,34	1000
>1000	2,85	2,75	2,66	2,58	2,52	2,46	2,40	2,35	2,31	>1000
df2/df1	11	12	13	14	15	16	17	18	19	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 2 Distribución F ($P = 0,001$) (Continuación)

Valores del estadístico F para $p = 0,001$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	20	22	24	26	28	30	35	40	45	df1\df2
3	126,42	126,16	125,94	125,75	125,59	125,45	125,17	124,96	124,80	3
4	46,10	45,92	45,77	45,64	45,53	45,43	45,24	45,09	44,98	4
5	25,40	25,25	25,13	25,03	24,95	24,87	24,72	24,60	24,51	5
6	17,12	17,00	16,90	16,81	16,74	16,67	16,54	16,45	16,37	6
7	12,93	12,82	12,73	12,66	12,59	12,53	12,41	12,33	12,26	7
8	10,48	10,38	10,30	10,22	10,16	10,11	10,00	9,92	9,86	8
9	8,90	8,80	8,72	8,66	8,60	8,55	8,45	8,37	8,31	9
10	7,80	7,71	7,64	7,57	7,52	7,47	7,37	7,30	7,24	10
11	7,01	6,92	6,85	6,79	6,73	6,68	6,59	6,52	6,46	11
12	6,41	6,32	6,25	6,19	6,14	6,09	6,00	5,93	5,87	12
13	5,93	5,85	5,78	5,72	5,67	5,63	5,54	5,47	5,41	13
14	5,56	5,48	5,41	5,35	5,30	5,25	5,17	5,10	5,05	14
15	5,25	5,17	5,10	5,04	4,99	4,95	4,86	4,80	4,74	15
16	4,99	4,91	4,85	4,79	4,74	4,70	4,61	4,55	4,49	16
17	4,78	4,70	4,63	4,58	4,53	4,48	4,40	4,33	4,28	17
18	4,59	4,51	4,45	4,39	4,34	4,30	4,22	4,15	4,10	18
19	4,43	4,35	4,29	4,23	4,19	4,14	4,06	3,99	3,94	19
20	4,29	4,21	4,15	4,09	4,05	4,01	3,92	3,86	3,81	20
22	4,06	3,98	3,92	3,86	3,82	3,78	3,69	3,63	3,58	22
24	3,87	3,80	3,74	3,68	3,63	3,59	3,51	3,45	3,40	24
26	3,72	3,65	3,59	3,53	3,49	3,45	3,36	3,30	3,25	26
28	3,60	3,52	3,46	3,41	3,36	3,32	3,24	3,18	3,13	28
30	3,49	3,42	3,36	3,30	3,26	3,22	3,14	3,07	3,02	30
35	3,29	3,22	3,16	3,10	3,06	3,02	2,93	2,87	2,82	35
40	3,15	3,07	3,01	2,96	2,91	2,87	2,79	2,73	2,68	40
45	3,04	2,96	2,90	2,85	2,80	2,76	2,68	2,62	2,57	45
50	2,95	2,88	2,82	2,77	2,72	2,68	2,60	2,53	2,48	50
60	2,83	2,76	2,69	2,64	2,60	2,56	2,47	2,41	2,36	60
70	2,74	2,67	2,61	2,56	2,51	2,47	2,39	2,32	2,27	70
80	2,68	2,61	2,55	2,49	2,45	2,41	2,32	2,26	2,21	80
100	2,59	2,52	2,46	2,41	2,36	2,32	2,24	2,17	2,12	100
200	2,42	2,35	2,29	2,24	2,19	2,15	2,07	2,00	1,95	200
500	2,33	2,26	2,20	2,14	2,10	2,05	1,97	1,90	1,85	500
1000	2,30	2,23	2,16	2,11	2,06	2,02	1,94	1,87	1,81	1000
>1000	2,27	2,20	2,14	2,08	2,04	1,99	1,91	1,84	1,78	>1000
df2\df1	20	22	24	26	28	30	35	40	45	df1\df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 2 Distribución F ($P = 0,001$) (Continuación)

Valores del estadístico F para $p = 0,001$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	50	60	70	80	100	200	500	1000	>1000	df1/df2
3	124,67	124,47	124,33	124,22	124,07	123,77	123,59	123,52	123,60	3
4	44,88	44,75	44,65	44,57	44,47	44,26	44,13	44,09	44,07	4
5	24,44	24,33	24,26	24,20	24,12	23,95	23,85	23,82	23,79	5
6	16,31	16,21	16,15	16,10	16,03	15,89	15,80	15,77	15,76	6
7	12,20	12,12	12,06	12,01	11,95	11,82	11,75	11,72	11,70	7
8	9,80	9,73	9,67	9,63	9,57	9,45	9,38	9,36	9,34	8
9	8,26	8,19	8,13	8,09	8,04	7,93	7,86	7,84	7,82	9
10	7,19	7,12	7,07	7,03	6,98	6,87	6,81	6,78	6,77	10
11	6,42	6,35	6,30	6,26	6,21	6,11	6,04	6,02	6,00	11
12	5,83	5,76	5,71	5,68	5,63	5,52	5,46	5,44	5,43	12
13	5,37	5,31	5,26	5,22	5,17	5,07	5,01	4,99	4,97	13
14	5,00	4,94	4,89	4,86	4,81	4,71	4,65	4,63	4,61	14
15	4,70	4,64	4,59	4,56	4,51	4,41	4,35	4,33	4,31	15
16	4,45	4,39	4,34	4,31	4,26	4,16	4,10	4,08	4,06	16
17	4,24	4,18	4,13	4,10	4,05	3,95	3,89	3,87	3,85	17
18	4,06	4,00	3,95	3,92	3,87	3,77	3,71	3,69	3,67	18
19	3,90	3,84	3,80	3,76	3,71	3,62	3,56	3,53	3,52	19
20	3,77	3,70	3,66	3,62	3,58	3,48	3,42	3,40	3,38	20
22	3,54	3,48	3,43	3,40	3,35	3,25	3,19	3,17	3,15	22
24	3,36	3,30	3,25	3,22	3,17	3,07	3,01	2,99	2,97	24
26	3,21	3,15	3,10	3,07	3,02	2,92	2,86	2,84	2,82	26
28	3,09	3,02	2,98	2,95	2,90	2,80	2,74	2,72	2,70	28
30	2,98	2,92	2,88	2,84	2,79	2,69	2,63	2,61	2,59	30
35	2,78	2,72	2,67	2,64	2,59	2,49	2,43	2,41	2,39	35
40	2,64	2,57	2,53	2,49	2,44	2,34	2,28	2,26	2,24	40
45	2,53	2,46	2,42	2,38	2,33	2,23	2,16	2,14	2,12	45
50	2,44	2,38	2,33	2,30	2,25	2,14	2,07	2,05	2,03	50
60	2,32	2,25	2,21	2,17	2,12	2,01	1,94	1,92	1,89	60
70	2,23	2,16	2,12	2,08	2,03	1,92	1,84	1,82	1,80	70
80	2,16	2,10	2,05	2,01	1,96	1,85	1,77	1,75	1,72	80
100	2,08	2,01	1,96	1,92	1,87	1,75	1,67	1,64	1,62	100
200	1,90	1,83	1,78	1,74	1,68	1,55	1,46	1,43	1,39	200
500	1,80	1,73	1,68	1,63	1,57	1,43	1,32	1,28	1,23	500
1000	1,77	1,70	1,64	1,60	1,53	1,38	1,27	1,22	1,16	1000
>1000	1,74	1,66	1,61	1,56	1,50	1,34	1,21	1,15	1,06	>1000
df2\df1	50	60	70	80	100	200	500	1000	>1000	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 3 Distribución F ($P = 0,01$)

Valores del estadístico F para $p = 0,01$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	1	2	3	4	5	6	7	8	9	10	df1\df2
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	3
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	4
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	5
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	6
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	7
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	8
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	9
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	10
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	11
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	12
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	13
14	8,86	6,51	5,56	5,04	4,70	4,46	4,28	4,14	4,03	3,94	14
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	15
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	16
17	8,40	6,11	5,19	4,67	4,34	4,10	3,93	3,79	3,68	3,59	17
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	18
19	8,19	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	19
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	20
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	22
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	24
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	26
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	28
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	30
35	7,42	5,27	4,40	3,91	3,59	3,37	3,20	3,07	2,96	2,88	35
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	40
45	7,23	5,11	4,25	3,77	3,45	3,23	3,07	2,94	2,83	2,74	45
50	7,17	5,06	4,20	3,72	3,41	3,19	3,02	2,89	2,79	2,70	50
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	60
70	7,01	4,92	4,07	3,60	3,29	3,07	2,91	2,78	2,67	2,59	70
80	6,96	4,88	4,04	3,56	3,26	3,04	2,87	2,74	2,64	2,55	80
100	6,90	4,82	3,98	3,51	3,21	2,99	2,82	2,69	2,59	2,50	100
200	6,76	4,71	3,88	3,41	3,11	2,89	2,73	2,60	2,50	2,41	200
500	6,69	4,65	3,82	3,36	3,05	2,84	2,68	2,55	2,44	2,36	500
1000	6,66	4,63	3,80	3,34	3,04	2,82	2,66	2,53	2,43	2,34	1000
>1000	1,04	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	>1000
df2\df1	1	2	3	4	5	6	7	8	9	10	df1\df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 3 Distribución F ($P = 0,01$) (Continuación)

Valores del estadístico F para $p = 0,01$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	11	12	13	14	15	16	17	18	19	df1\df2
3	27,13	27,05	26,98	26,92	26,87	26,83	26,79	26,75	26,72	3
4	14,45	14,37	14,31	14,25	14,20	14,15	14,11	14,08	14,05	4
5	9,96	9,89	9,82	9,77	9,72	9,68	9,64	9,61	9,58	5
6	7,79	7,72	7,66	7,61	7,56	7,52	7,48	7,45	7,42	6
7	6,54	6,47	6,41	6,36	6,31	6,28	6,24	6,21	6,18	7
8	5,73	5,67	5,61	5,56	5,52	5,48	5,44	5,41	5,38	8
9	5,18	5,11	5,05	5,01	4,96	4,92	4,89	4,86	4,83	9
10	4,77	4,71	4,65	4,60	4,56	4,52	4,49	4,46	4,43	10
11	4,46	4,40	4,34	4,29	4,25	4,21	4,18	4,15	4,12	11
12	4,22	4,16	4,10	4,05	4,01	3,97	3,94	3,91	3,88	12
13	4,02	3,96	3,91	3,86	3,82	3,78	3,75	3,72	3,69	13
14	3,86	3,80	3,75	3,70	3,66	3,62	3,59	3,56	3,53	14
15	3,73	3,67	3,61	3,56	3,52	3,49	3,45	3,42	3,40	15
16	3,62	3,55	3,50	3,45	3,41	3,37	3,34	3,31	3,28	16
17	3,52	3,46	3,40	3,35	3,31	3,27	3,24	3,21	3,19	17
18	3,43	3,37	3,32	3,27	3,23	3,19	3,16	3,13	3,10	18
19	3,36	3,30	3,24	3,19	3,15	3,12	3,08	3,05	3,03	19
20	3,29	3,23	3,18	3,13	3,09	3,05	3,02	2,99	2,96	20
22	3,18	3,12	3,07	3,02	2,98	2,94	2,91	2,88	2,85	22
24	3,09	3,03	2,98	2,93	2,89	2,85	2,82	2,79	2,76	24
26	3,02	2,96	2,90	2,86	2,82	2,78	2,75	2,72	2,69	26
28	2,96	2,90	2,84	2,79	2,75	2,72	2,68	2,65	2,63	28
30	2,91	2,84	2,79	2,74	2,70	2,66	2,63	2,60	2,57	30
35	2,80	2,74	2,69	2,64	2,60	2,56	2,53	2,50	2,47	35
40	2,73	2,66	2,61	2,56	2,52	2,48	2,45	2,42	2,39	40
45	2,67	2,61	2,55	2,51	2,46	2,43	2,39	2,36	2,34	45
50	2,63	2,56	2,51	2,46	2,42	2,38	2,35	2,32	2,29	50
60	2,56	2,50	2,44	2,39	2,35	2,31	2,28	2,25	2,22	60
70	2,51	2,45	2,40	2,35	2,31	2,27	2,23	2,20	2,18	70
80	2,48	2,42	2,36	2,31	2,27	2,23	2,20	2,17	2,14	80
100	2,43	2,37	2,31	2,27	2,22	2,19	2,15	2,12	2,09	100
200	2,34	2,27	2,22	2,17	2,13	2,09	2,06	2,03	2,00	200
500	2,28	2,22	2,17	2,12	2,07	2,04	2,00	1,97	1,94	500
1000	2,27	2,20	2,15	2,10	2,06	2,02	1,98	1,95	1,92	1000
>1000	2,25	2,19	2,13	2,08	2,04	2,00	1,97	1,94	1,91	>1000
df2\df1	11	12	13	14	15	16	17	18	19	df1\df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 3 Distribución F ($P = 0,01$) (Continuación)

Valores del estadístico F para $p = 0,01$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación ($df1$) y de la fuente de error ($df2$) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

$df2 \backslash df1$	20	22	24	26	28	30	35	40	45	$df1 \backslash df2$
3	26,69	26,64	26,60	26,56	26,53	26,50	26,45	26,41	26,38	3
4	14,02	13,97	13,93	13,89	13,86	13,84	13,79	13,75	13,71	4
5	9,55	9,51	9,47	9,43	9,40	9,38	9,33	9,29	9,26	5
6	7,40	7,35	7,31	7,28	7,25	7,23	7,18	7,14	7,11	6
7	6,16	6,11	6,07	6,04	6,02	5,99	5,94	5,91	5,88	7
8	5,36	5,32	5,28	5,25	5,22	5,20	5,15	5,12	5,09	8
9	4,81	4,77	4,73	4,70	4,67	4,65	4,60	4,57	4,54	9
10	4,41	4,36	4,33	4,30	4,27	4,25	4,20	4,17	4,14	10
11	4,10	4,06	4,02	3,99	3,96	3,94	3,89	3,86	3,83	11
12	3,86	3,82	3,78	3,75	3,72	3,70	3,65	3,62	3,59	12
13	3,66	3,62	3,59	3,56	3,53	3,51	3,46	3,43	3,40	13
14	3,51	3,46	3,43	3,40	3,37	3,35	3,30	3,27	3,24	14
15	3,37	3,33	3,29	3,26	3,24	3,21	3,17	3,13	3,10	15
16	3,26	3,22	3,18	3,15	3,12	3,10	3,05	3,02	2,99	16
17	3,16	3,12	3,08	3,05	3,03	3,00	2,96	2,92	2,89	17
18	3,08	3,03	3,00	2,97	2,94	2,92	2,87	2,84	2,81	18
19	3,00	2,96	2,92	2,89	2,87	2,84	2,80	2,76	2,73	19
20	2,94	2,90	2,86	2,83	2,80	2,78	2,73	2,69	2,67	20
22	2,83	2,78	2,75	2,72	2,69	2,67	2,62	2,58	2,55	22
24	2,74	2,70	2,66	2,63	2,60	2,58	2,53	2,49	2,46	24
26	2,66	2,62	2,58	2,55	2,53	2,50	2,45	2,42	2,39	26
28	2,60	2,56	2,52	2,49	2,46	2,44	2,39	2,35	2,32	28
30	2,55	2,51	2,47	2,44	2,41	2,39	2,34	2,30	2,27	30
35	2,44	2,40	2,36	2,33	2,31	2,28	2,23	2,19	2,16	35
40	2,37	2,33	2,29	2,26	2,23	2,20	2,15	2,11	2,08	40
45	2,31	2,27	2,23	2,20	2,17	2,14	2,09	2,05	2,02	45
50	2,27	2,22	2,18	2,15	2,12	2,10	2,05	2,01	1,97	50
60	2,20	2,15	2,12	2,08	2,05	2,03	1,98	1,94	1,90	60
70	2,15	2,11	2,07	2,03	2,01	1,98	1,93	1,89	1,85	70
80	2,12	2,07	2,03	2,00	1,97	1,94	1,89	1,85	1,82	80
100	2,07	2,02	1,98	1,95	1,92	1,89	1,84	1,80	1,76	100
200	1,97	1,93	1,89	1,85	1,82	1,79	1,74	1,69	1,66	200
500	1,92	1,87	1,83	1,79	1,76	1,74	1,68	1,63	1,60	500
1000	1,90	1,85	1,81	1,77	1,74	1,72	1,66	1,61	1,58	1000
>1000	1,88	1,83	1,79	1,76	1,73	1,70	1,64	1,59	1,56	>1000
$df2 \backslash df1$	20	22	24	26	28	30	35	40	45	$df1 \backslash df2$

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 3 Distribución F ($P = 0,01$) (Continuación)

Valores del estadístico F para $p = 0,01$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	50	60	70	80	100	200	500	1000	>1000	df1/df2
3	26,35	26,32	26,29	26,27	26,24	26,18	26,15	26,13	26,15	3
4	13,69	13,65	13,63	13,61	13,58	13,52	13,49	13,47	13,47	4
5	9,24	9,20	9,18	9,16	9,13	9,08	9,04	9,03	9,02	5
6	7,09	7,06	7,03	7,01	6,99	6,93	6,90	6,89	6,89	6
7	5,86	5,82	5,80	5,78	5,75	5,70	5,67	5,66	5,65	7
8	5,07	5,03	5,01	4,99	4,96	4,91	4,88	4,87	4,86	8
9	4,52	4,48	4,46	4,44	4,42	4,36	4,33	4,32	4,32	9
10	4,12	4,08	4,06	4,04	4,01	3,96	3,93	3,92	3,91	10
11	3,81	3,78	3,75	3,73	3,71	3,66	3,62	3,61	3,60	11
12	3,57	3,54	3,51	3,49	3,47	3,41	3,38	3,37	3,36	12
13	3,38	3,34	3,32	3,30	3,27	3,22	3,19	3,18	3,17	13
14	3,22	3,18	3,16	3,14	3,11	3,06	3,03	3,01	3,01	14
15	3,08	3,05	3,02	3,00	2,98	2,92	2,89	2,88	2,87	15
16	2,97	2,93	2,91	2,89	2,86	2,81	2,78	2,76	2,75	16
17	2,87	2,83	2,81	2,79	2,76	2,71	2,68	2,66	2,65	17
18	2,78	2,75	2,72	2,71	2,68	2,62	2,59	2,58	2,57	18
19	2,71	2,67	2,65	2,63	2,60	2,55	2,51	2,50	2,49	19
20	2,64	2,61	2,58	2,56	2,54	2,48	2,44	2,43	2,42	20
22	2,53	2,50	2,47	2,45	2,42	2,36	2,33	2,32	2,31	22
24	2,44	2,40	2,38	2,36	2,33	2,27	2,24	2,22	2,21	24
26	2,36	2,33	2,30	2,28	2,25	2,19	2,16	2,14	2,13	26
28	2,30	2,26	2,24	2,22	2,19	2,13	2,09	2,08	2,07	28
30	2,25	2,21	2,18	2,16	2,13	2,07	2,03	2,02	2,01	30
35	2,14	2,10	2,07	2,05	2,02	1,96	1,92	1,90	1,89	35
40	2,06	2,02	1,99	1,97	1,94	1,87	1,83	1,82	1,81	40
45	2,00	1,96	1,93	1,91	1,88	1,81	1,77	1,75	1,74	45
50	1,95	1,91	1,88	1,86	1,82	1,76	1,71	1,70	1,69	50
60	1,88	1,84	1,81	1,78	1,75	1,68	1,63	1,62	1,60	60
70	1,83	1,78	1,75	1,73	1,70	1,62	1,57	1,56	1,54	70
80	1,79	1,75	1,71	1,69	1,65	1,58	1,53	1,51	1,50	80
100	1,74	1,69	1,66	1,63	1,60	1,52	1,47	1,45	1,43	100
200	1,63	1,58	1,55	1,52	1,48	1,39	1,33	1,30	1,28	200
500	1,57	1,52	1,48	1,45	1,41	1,31	1,23	1,20	1,17	500
1000	1,54	1,50	1,46	1,43	1,38	1,28	1,19	1,16	1,12	1000
>1000	1,53	1,48	1,44	1,41	1,36	1,25	1,16	1,11	1,05	>1000
df2/df1	50	60	70	80	100	200	500	1000	>1000	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 4 Distribución F ($P = 0,05$)

Valores del estadístico F para $p = 0,05$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	1	2	3	4	5	6	7	8	9	10	df1\df2
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	3
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	4
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	5
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	6
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	7
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	8
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	9
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	10
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	11
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	12
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	13
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	14
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	15
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	16
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	17
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	18
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	19
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	20
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	22
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	24
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	26
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	28
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	30
35	4,12	3,27	2,87	2,64	2,49	2,37	2,29	2,22	2,16	2,11	35
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	40
45	4,06	3,20	2,81	2,58	2,42	2,31	2,22	2,15	2,10	2,05	45
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	50
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	60
70	3,98	3,13	2,74	2,50	2,35	2,23	2,14	2,07	2,02	1,97	70
80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2,00	1,95	80
100	3,94	3,09	2,70	2,46	2,31	2,19	2,10	2,03	1,97	1,93	100
200	3,89	3,04	2,65	2,42	2,26	2,14	2,06	1,98	1,93	1,88	200
500	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	500
1000	3,85	3,00	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1000
>1000	1,04	3,00	2,61	2,37	2,21	2,10	2,01	1,94	1,88	1,83	>1000
df2\df1	1	2	3	4	5	6	7	8	9	10	df1\df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 4 Distribución F ($P = 0,05$) (Continuación)

Valores del estadístico F para $p = 0,05$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	11	12	13	14	15	16	17	18	19	df1/df2
3	8,76	8,74	8,73	8,71	8,70	8,69	8,68	8,67	8,67	3
4	5,94	5,91	5,89	5,87	5,86	5,84	5,83	5,82	5,81	4
5	4,70	4,68	4,66	4,64	4,62	4,60	4,59	4,58	4,57	5
6	4,03	4,00	3,98	3,96	3,94	3,92	3,91	3,90	3,88	6
7	3,60	3,57	3,55	3,53	3,51	3,49	3,48	3,47	3,46	7
8	3,31	3,28	3,26	3,24	3,22	3,20	3,19	3,17	3,16	8
9	3,10	3,07	3,05	3,03	3,01	2,99	2,97	2,96	2,95	9
10	2,94	2,91	2,89	2,86	2,85	2,83	2,81	2,80	2,79	10
11	2,82	2,79	2,76	2,74	2,72	2,70	2,69	2,67	2,66	11
12	2,72	2,69	2,66	2,64	2,62	2,60	2,58	2,57	2,56	12
13	2,63	2,60	2,58	2,55	2,53	2,51	2,50	2,48	2,47	13
14	2,57	2,53	2,51	2,48	2,46	2,44	2,43	2,41	2,40	14
15	2,51	2,48	2,45	2,42	2,40	2,38	2,37	2,35	2,34	15
16	2,46	2,42	2,40	2,37	2,35	2,33	2,32	2,30	2,29	16
17	2,41	2,38	2,35	2,33	2,31	2,29	2,27	2,26	2,24	17
18	2,37	2,34	2,31	2,29	2,27	2,25	2,23	2,22	2,20	18
19	2,34	2,31	2,28	2,26	2,23	2,21	2,20	2,18	2,17	19
20	2,31	2,28	2,25	2,23	2,20	2,18	2,17	2,15	2,14	20
22	2,26	2,23	2,20	2,17	2,15	2,13	2,11	2,10	2,08	22
24	2,22	2,18	2,15	2,13	2,11	2,09	2,07	2,05	2,04	24
26	2,18	2,15	2,12	2,09	2,07	2,05	2,03	2,02	2,00	26
28	2,15	2,12	2,09	2,06	2,04	2,02	2,00	1,99	1,97	28
30	2,13	2,09	2,06	2,04	2,01	1,99	1,98	1,96	1,95	30
35	2,08	2,04	2,01	1,99	1,96	1,94	1,92	1,91	1,89	35
40	2,04	2,00	1,97	1,95	1,92	1,90	1,89	1,87	1,85	40
45	2,01	1,97	1,94	1,92	1,89	1,87	1,86	1,84	1,82	45
50	1,99	1,95	1,92	1,89	1,87	1,85	1,83	1,81	1,80	50
60	1,95	1,92	1,89	1,86	1,84	1,82	1,80	1,78	1,76	60
70	1,93	1,89	1,86	1,84	1,81	1,79	1,77	1,75	1,74	70
80	1,91	1,88	1,84	1,82	1,79	1,77	1,75	1,73	1,72	80
100	1,89	1,85	1,82	1,79	1,77	1,75	1,73	1,71	1,69	100
200	1,84	1,80	1,77	1,74	1,72	1,69	1,67	1,66	1,64	200
500	1,81	1,77	1,74	1,71	1,69	1,66	1,64	1,62	1,61	500
1000	1,80	1,76	1,73	1,70	1,68	1,65	1,63	1,61	1,60	1000
>1000	1,79	1,75	1,72	1,69	1,67	1,64	1,62	1,61	1,59	>1000
df2/df1	11	12	13	14	15	16	17	18	19	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 4 Distribución F ($P = 0,05$) (Continuación)

Valores del estadístico F para $p = 0,05$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	20	22	24	26	28	30	35	40	45	df1/df2
3	8,66	8,65	8,64	8,63	8,62	8,62	8,60	8,59	8,59	3
4	5,80	5,79	5,77	5,76	5,75	5,75	5,73	5,72	5,71	4
5	4,56	4,54	4,53	4,52	4,50	4,50	4,48	4,46	4,45	5
6	3,87	3,86	3,84	3,83	3,82	3,81	3,79	3,77	3,76	6
7	3,44	3,43	3,41	3,40	3,39	3,38	3,36	3,34	3,33	7
8	3,15	3,13	3,12	3,10	3,09	3,08	3,06	3,04	3,03	8
9	2,94	2,92	2,90	2,89	2,87	2,86	2,84	2,83	2,81	9
10	2,77	2,75	2,74	2,72	2,71	2,70	2,68	2,66	2,65	10
11	2,65	2,63	2,61	2,59	2,58	2,57	2,55	2,53	2,52	11
12	2,54	2,52	2,51	2,49	2,48	2,47	2,44	2,43	2,41	12
13	2,46	2,44	2,42	2,41	2,39	2,38	2,36	2,34	2,33	13
14	2,39	2,37	2,35	2,33	2,32	2,31	2,28	2,27	2,25	14
15	2,33	2,31	2,29	2,27	2,26	2,25	2,22	2,20	2,19	15
16	2,28	2,25	2,24	2,22	2,21	2,19	2,17	2,15	2,14	16
17	2,23	2,21	2,19	2,17	2,16	2,15	2,12	2,10	2,09	17
18	2,19	2,17	2,15	2,13	2,12	2,11	2,08	2,06	2,05	18
19	2,16	2,13	2,11	2,10	2,08	2,07	2,05	2,03	2,01	19
20	2,12	2,10	2,08	2,07	2,05	2,04	2,01	1,99	1,98	20
22	2,07	2,05	2,03	2,01	2,00	1,98	1,96	1,94	1,92	22
24	2,03	2,00	1,98	1,97	1,95	1,94	1,91	1,89	1,88	24
26	1,99	1,97	1,95	1,93	1,91	1,90	1,87	1,85	1,84	26
28	1,96	1,93	1,91	1,90	1,88	1,87	1,84	1,82	1,80	28
30	1,93	1,91	1,89	1,87	1,85	1,84	1,81	1,79	1,77	30
35	1,88	1,85	1,83	1,82	1,80	1,79	1,76	1,74	1,72	35
40	1,84	1,81	1,79	1,77	1,76	1,74	1,72	1,69	1,67	40
45	1,81	1,78	1,76	1,74	1,73	1,71	1,68	1,66	1,64	45
50	1,78	1,76	1,74	1,72	1,70	1,69	1,66	1,63	1,61	50
60	1,75	1,72	1,70	1,68	1,66	1,65	1,62	1,59	1,57	60
70	1,72	1,70	1,67	1,65	1,64	1,62	1,59	1,57	1,55	70
80	1,70	1,68	1,65	1,63	1,62	1,60	1,57	1,54	1,52	80
100	1,68	1,65	1,63	1,61	1,59	1,57	1,54	1,52	1,49	100
200	1,62	1,60	1,57	1,55	1,53	1,52	1,48	1,46	1,43	200
500	1,59	1,56	1,54	1,52	1,50	1,48	1,45	1,42	1,40	500
1000	1,58	1,55	1,53	1,51	1,49	1,47	1,43	1,41	1,38	1000
>1000	1,57	1,54	1,52	1,50	1,48	1,46	1,42	1,40	1,37	>1000
df2\df1	20	22	24	26	28	30	35	40	45	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

TABLA 4 Distribución F ($P = 0,05$) (Continuación)

Valores del estadístico F para $p = 0,05$. El valor del estadístico F se encuentra en el cuerpo de la tabla, en función de los grados de libertad de la fuente de variación (df1) y de la fuente de error (df2) (filas superior o inferior y columnas izquierda o derecha, respectivamente).

df2\df1	50	60	70	80	100	200	500	1000	>1000	df1/df2
3	8,58	8,57	8,57	8,56	8,55	8,54	8,53	8,53	8,54	3
4	5,70	5,69	5,68	5,67	5,66	5,65	5,64	5,63	5,63	4
5	4,44	4,43	4,42	4,42	4,41	4,39	4,37	4,37	4,36	5
6	3,75	3,74	3,73	3,72	3,71	3,69	3,68	3,67	3,67	6
7	3,32	3,30	3,29	3,29	3,27	3,25	3,24	3,23	3,23	7
8	3,02	3,01	2,99	2,99	2,97	2,95	2,94	2,93	2,93	8
9	2,80	2,79	2,78	2,77	2,76	2,73	2,72	2,71	2,71	9
10	2,64	2,62	2,61	2,60	2,59	2,56	2,55	2,54	2,54	10
11	2,51	2,49	2,48	2,47	2,46	2,43	2,42	2,41	2,41	11
12	2,40	2,38	2,37	2,36	2,35	2,32	2,31	2,30	2,30	12
13	2,31	2,30	2,28	2,27	2,26	2,23	2,22	2,21	2,21	13
14	2,24	2,22	2,21	2,20	2,19	2,16	2,14	2,14	2,13	14
15	2,18	2,16	2,15	2,14	2,12	2,10	2,08	2,07	2,07	15
16	2,12	2,11	2,09	2,08	2,07	2,04	2,02	2,02	2,01	16
17	2,08	2,06	2,05	2,03	2,02	1,99	1,97	1,97	1,96	17
18	2,04	2,02	2,00	1,99	1,98	1,95	1,93	1,92	1,92	18
19	2,00	1,98	1,97	1,96	1,94	1,91	1,89	1,88	1,88	19
20	1,97	1,95	1,93	1,92	1,91	1,88	1,86	1,85	1,84	20
22	1,91	1,89	1,88	1,86	1,85	1,82	1,80	1,79	1,78	22
24	1,86	1,84	1,83	1,82	1,80	1,77	1,75	1,74	1,73	24
26	1,82	1,80	1,79	1,78	1,76	1,73	1,71	1,70	1,69	26
28	1,79	1,77	1,75	1,74	1,73	1,69	1,67	1,66	1,66	28
30	1,76	1,74	1,72	1,71	1,70	1,66	1,64	1,63	1,62	30
35	1,70	1,68	1,66	1,65	1,63	1,60	1,57	1,57	1,56	35
40	1,66	1,64	1,62	1,61	1,59	1,55	1,53	1,52	1,51	40
45	1,63	1,60	1,59	1,57	1,55	1,51	1,49	1,48	1,47	45
50	1,60	1,58	1,56	1,54	1,52	1,48	1,46	1,45	1,44	50
60	1,56	1,53	1,52	1,50	1,48	1,44	1,41	1,40	1,39	60
70	1,53	1,50	1,49	1,47	1,45	1,40	1,37	1,36	1,35	70
80	1,51	1,48	1,46	1,45	1,43	1,38	1,35	1,34	1,33	80
100	1,48	1,45	1,43	1,41	1,39	1,34	1,31	1,30	1,28	100
200	1,41	1,39	1,36	1,35	1,32	1,26	1,22	1,21	1,19	200
500	1,38	1,35	1,32	1,30	1,28	1,21	1,16	1,14	1,12	500
1000	1,36	1,33	1,31	1,29	1,26	1,19	1,13	1,11	1,08	1000
>1000	1,35	1,32	1,30	1,28	1,25	1,17	1,11	1,08	1,03	>1000
df2\df1	50	60	70	80	100	200	500	1000	>1000	df1/df2

FUENTE: <http://thesaurus.maths.org/dictionary/map/word/647>. Reproducida con la autorización del autor.

**TABLA 5 Valores críticos de la distribución de Rango Studentizado
para grados de libertad de error entre 1 y 10**

Error df	Alpha	k = number of means or number of steps between ordered means							Alpha	Error df
		2	3	4	5	6	7	8		
1	0,100	8,951	13,453	16,378	18,504	20,164	21,516	22,649	0,100	1
	0,050	18,066	27,066	32,925	37,149	40,481	43,203	45,501	0,050	
	0,010	93,157	138,306	168,728	189,173	206,203	219,531	231,719	0,010	
2	0,100	4,136	5,736	6,777	7,540	8,142	8,635	9,052	0,100	2
	0,050	6,101	8,344	9,813	10,891	11,744	12,444	13,039	0,050	
	0,010	14,250	19,206	22,522	24,897	26,813	28,382	29,750	0,010	
3	0,100	3,331	4,469	5,200	5,739	6,163	6,511	6,807	0,100	3
	0,050	4,508	5,914	6,828	7,504	8,039	8,480	8,855	0,050	
	0,010	8,314	10,664	12,225	13,362	14,284	15,032	15,691	0,010	
4	0,100	3,017	3,977	4,587	5,036	5,389	5,680	5,926	0,100	4
	0,050	3,932	5,044	5,761	6,290	6,709	7,055	7,349	0,050	
	0,010	6,541	8,152	9,211	9,988	10,613	11,127	11,573	0,010	
5	0,100	2,852	3,719	4,265	4,665	4,980	5,239	5,459	0,100	5
	0,050	3,639	4,605	5,221	5,676	6,035	6,332	6,585	0,050	
	0,010	5,727	7,002	7,828	8,442	8,933	9,339	9,691	0,010	
6	0,100	2,750	3,560	4,066	4,436	4,727	4,966	5,169	0,100	6
	0,050	3,464	4,342	4,898	5,307	5,630	5,897	6,124	0,050	
	0,010	5,268	6,351	7,050	7,572	7,988	8,337	8,630	0,010	
7	0,100	2,681	3,452	3,932	4,281	4,556	4,781	4,972	0,100	7
	0,050	3,347	4,167	4,683	5,062	5,361	5,607	5,817	0,050	
	0,010	4,967	5,934	6,557	7,018	7,386	7,692	7,953	0,010	
8	0,100	2,631	3,375	3,835	4,169	4,432	4,647	4,829	0,100	8
	0,050	3,264	4,043	4,531	4,888	5,169	5,400	5,598	0,050	
	0,010	4,761	5,648	6,219	6,637	6,970	7,248	7,485	0,010	
9	0,100	2,594	3,317	3,762	4,085	4,338	4,546	4,721	0,100	9
	0,050	3,202	3,951	4,416	4,757	5,025	5,246	5,433	0,050	
	0,010	4,609	5,439	5,969	6,358	6,666	6,924	7,145	0,010	
10	0,100	2,564	3,271	3,705	4,019	4,264	4,466	4,636	0,100	10
	0,050	3,153	3,879	4,328	4,656	4,913	5,126	5,305	0,050	
	0,010	4,495	5,282	5,780	6,145	6,435	6,677	6,884	0,010	
Error df	Alpha	2	3	4	5	6	7	8	Alpha	Error df
		k = number of means or number of steps between ordered means								

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 5 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error entre 1 y 10 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		9	10	11	12	13	14		
1	0,100	23,627	24,478	25,239	25,918	26,520	27,081	0,100	1
	0,050	47,482	49,220	50,780	52,161	53,346	54,469	0,050	
	0,010	241,881	250,842	258,985	266,339	271,083	277,480	0,010	
2	0,100	9,412	9,728	10,011	10,264	10,491	10,701	0,100	2
	0,050	13,552	14,003	14,407	14,761	15,086	15,386	0,050	
	0,010	30,923	31,929	32,874	33,644	34,373	35,059	0,010	
3	0,100	7,063	7,287	7,488	7,669	7,832	7,982	0,100	3
	0,050	9,180	9,465	9,721	9,948	10,156	10,347	0,050	
	0,010	16,254	16,752	17,197	17,569	17,926	18,260	0,010	
4	0,100	6,140	6,328	6,496	6,647	6,784	6,909	0,100	4
	0,050	7,604	7,829	8,031	8,212	8,376	8,527	0,050	
	0,010	11,960	12,301	12,613	12,871	13,117	13,350	0,010	
5	0,100	5,649	5,817	5,967	6,101	6,224	6,337	0,100	5
	0,050	6,804	6,997	7,171	7,325	7,467	7,598	0,050	
	0,010	9,997	10,265	10,511	10,718	10,916	11,098	0,010	
6	0,100	5,345	5,499	5,638	5,762	5,876	5,980	0,100	6
	0,050	6,321	6,495	6,651	6,791	6,918	7,036	0,050	
	0,010	8,887	9,115	9,325	9,500	9,668	9,824	0,010	
7	0,100	5,137	5,283	5,414	5,531	5,638	5,736	0,100	7
	0,050	5,999	6,160	6,304	6,433	6,551	6,660	0,050	
	0,010	8,180	8,383	8,567	8,723	8,872	9,009	0,010	
8	0,100	4,987	5,126	5,251	5,363	5,465	5,558	0,100	8
	0,050	5,769	5,920	6,055	6,177	6,288	6,390	0,050	
	0,010	7,693	7,876	8,043	8,185	8,321	8,446	0,010	
9	0,100	4,873	5,007	5,127	5,235	5,333	5,423	0,100	9
	0,050	5,596	5,740	5,868	5,985	6,090	6,187	0,050	
	0,010	7,336	7,506	7,660	7,793	7,918	8,033	0,010	
10	0,100	4,784	4,914	5,030	5,134	5,230	5,317	0,100	10
	0,050	5,462	5,600	5,723	5,835	5,936	6,029	0,050	
	0,010	7,064	7,223	7,368	7,493	7,610	7,719	0,010	
Error df	Alpha	9	10	11	12	13	14	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 5 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error entre 1 y 10 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		15	16	17	18	19	20		
1	0,100	27,606	28,087	28,530	28,950	29,346	29,720	0,100	1
	0,050	55,530	56,486	57,349	58,172	58,941	59,663	0,050	
	0,010	283,748	289,019	293,348	298,008	302,417	306,636	0,010	
2	0,100	10,894	11,074	11,240	11,396	11,542	11,679	0,100	2
	0,050	15,662	15,921	16,157	16,379	16,588	16,784	0,050	
	0,010	35,693	36,321	36,804	37,316	37,798	38,253	0,010	
3	0,100	8,120	8,249	8,369	8,480	8,585	8,684	0,100	3
	0,050	10,524	10,689	10,841	10,985	11,119	11,245	0,050	
	0,010	18,569	18,868	19,124	19,374	19,610	19,832	0,010	
4	0,100	7,025	7,133	7,234	7,328	7,416	7,499	0,100	4
	0,050	8,666	8,796	8,918	9,031	9,137	9,238	0,050	
	0,010	13,563	13,766	13,952	14,125	14,288	14,442	0,010	
5	0,100	6,440	6,536	6,626	6,710	6,789	6,864	0,100	5
	0,050	7,718	7,830	7,935	8,033	8,125	8,211	0,050	
	0,010	11,267	11,425	11,577	11,711	11,841	11,964	0,010	
6	0,100	6,076	6,165	6,248	6,325	6,398	6,467	0,100	6
	0,050	7,145	7,245	7,340	7,428	7,511	7,589	0,050	
	0,010	9,967	10,101	10,230	10,346	10,456	10,561	0,010	
7	0,100	5,826	5,910	5,989	6,062	6,131	6,196	0,100	7
	0,050	6,760	6,853	6,941	7,022	7,099	7,171	0,050	
	0,010	9,137	9,255	9,369	9,473	9,571	9,664	0,010	
8	0,100	5,645	5,725	5,800	5,870	5,935	5,997	0,100	8
	0,050	6,484	6,572	6,654	6,731	6,803	6,871	0,050	
	0,010	8,562	8,669	8,773	8,868	8,957	9,042	0,010	
9	0,100	5,506	5,583	5,656	5,723	5,786	5,846	0,100	9
	0,050	6,277	6,360	6,439	6,512	6,580	6,645	0,050	
	0,010	8,140	8,240	8,337	8,425	8,508	8,586	0,010	
10	0,100	5,398	5,472	5,542	5,608	5,669	5,727	0,100	10
	0,050	6,115	6,195	6,270	6,340	6,406	6,468	0,050	
	0,010	7,820	7,914	8,004	8,086	8,164	8,237	0,010	
Error df	Alpha	15	16	17	18	19	20	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 6 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error 11 y 20

Error df	Alpha	k = number of means or number of steps between ordered means							Alpha	Error df
		2	3	4	5	6	7	8		
11	0,100	2,541	3,235	3,659	3,965	4,205	4,402	4,568	0,100	11
	0,050	3,115	3,822	4,258	4,575	4,824	5,030	5,203	0,050	
	0,010	4,405	5,157	5,631	5,979	6,254	6,484	6,679	0,010	
12	0,100	2,522	3,205	3,622	3,922	4,157	4,349	4,512	0,100	12
	0,050	3,083	3,775	4,200	4,509	4,752	4,951	5,120	0,050	
	0,010	4,333	5,056	5,511	5,844	6,108	6,328	6,515	0,010	
13	0,100	2,506	3,180	3,590	3,885	4,116	4,305	4,465	0,100	13
	0,050	3,057	3,736	4,152	4,454	4,691	4,885	5,050	0,050	
	0,010	4,272	4,973	5,412	5,733	5,987	6,199	6,379	0,010	
14	0,100	2,492	3,158	3,563	3,855	4,082	4,268	4,425	0,100	14
	0,050	3,035	3,703	4,112	4,408	4,640	4,830	4,992	0,050	
	0,010	4,221	4,903	5,330	5,642	5,886	6,092	6,265	0,010	
15	0,100	2,480	3,140	3,541	3,828	4,052	4,235	4,390	0,100	15
	0,050	3,016	3,675	4,077	4,368	4,596	4,783	4,941	0,050	
	0,010	4,178	4,844	5,259	5,563	5,802	6,000	6,168	0,010	
16	0,100	2,470	3,125	3,521	3,805	4,026	4,207	4,360	0,100	16
	0,050	3,000	3,651	4,047	4,334	4,558	4,742	4,897	0,050	
	0,010	4,141	4,793	5,199	5,496	5,728	5,922	6,085	0,010	
17	0,100	2,461	3,111	3,503	3,785	4,004	4,183	4,334	0,100	17
	0,050	2,985	3,630	4,021	4,304	4,525	4,706	4,859	0,050	
	0,010	4,109	4,749	5,147	5,437	5,664	5,853	6,013	0,010	
18	0,100	2,453	3,098	3,488	3,767	3,984	4,161	4,311	0,100	18
	0,050	2,973	3,611	3,998	4,277	4,495	4,674	4,825	0,050	
	0,010	4,081	4,711	5,101	5,386	5,609	5,794	5,950	0,010	
19	0,100	2,446	3,088	3,474	3,751	3,966	4,142	4,290	0,100	19
	0,050	2,962	3,594	3,978	4,254	4,470	4,646	4,795	0,050	
	0,010	4,056	4,677	5,060	5,341	5,559	5,741	5,894	0,010	
20	0,100	2,440	3,078	3,462	3,737	3,950	4,125	4,272	0,100	20
	0,050	2,952	3,579	3,959	4,233	4,446	4,621	4,769	0,050	
	0,010	4,034	4,646	5,024	5,300	5,515	5,693	5,844	0,010	
Error df	Alpha	2	3	4	5	6	7	8	Alpha	Error df
		k = number of means or number of steps between ordered means								

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/216.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 6 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error 11 y 20 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		9	10	11	12	13	14		
11	0,100	4,712	4,838	4,951	5,053	5,146	5,231	0,100	11
	0,050	5,354	5,487	5,607	5,714	5,812	5,902	0,050	
	0,010	6,850	7,001	7,138	7,257	7,369	7,471	0,010	
12	0,100	4,652	4,776	4,887	4,986	5,077	5,160	0,100	12
	0,050	5,266	5,396	5,511	5,616	5,711	5,798	0,050	
	0,010	6,677	6,822	6,953	7,066	7,173	7,271	0,010	
13	0,100	4,603	4,724	4,832	4,930	5,019	5,101	0,100	13
	0,050	5,193	5,319	5,432	5,534	5,627	5,711	0,050	
	0,010	6,535	6,674	6,800	6,909	7,012	7,107	0,010	
14	0,100	4,560	4,680	4,786	4,882	4,970	5,050	0,100	14
	0,050	5,131	5,254	5,365	5,464	5,555	5,638	0,050	
	0,010	6,416	6,551	6,672	6,778	6,877	6,968	0,010	
15	0,100	4,524	4,642	4,747	4,841	4,927	5,006	0,100	15
	0,050	5,078	5,199	5,307	5,404	5,493	5,575	0,050	
	0,010	6,316	6,446	6,563	6,666	6,762	6,850	0,010	
16	0,100	4,492	4,609	4,712	4,805	4,890	4,968	0,100	16
	0,050	5,032	5,151	5,257	5,353	5,440	5,520	0,050	
	0,010	6,229	6,355	6,469	6,571	6,663	6,749	0,010	
17	0,100	4,465	4,579	4,682	4,774	4,858	4,935	0,100	17
	0,050	4,992	5,109	5,213	5,307	5,393	5,472	0,050	
	0,010	6,153	6,276	6,388	6,487	6,577	6,661	0,010	
18	0,100	4,440	4,554	4,655	4,746	4,829	4,905	0,100	18
	0,050	4,956	5,071	5,174	5,267	5,352	5,430	0,050	
	0,010	6,087	6,207	6,316	6,413	6,501	6,583	0,010	
19	0,100	4,418	4,531	4,631	4,721	4,803	4,879	0,100	19
	0,050	4,925	5,038	5,140	5,232	5,315	5,392	0,050	
	0,010	6,028	6,146	6,253	6,348	6,434	6,515	0,010	
20	0,100	4,398	4,510	4,609	4,699	4,780	4,855	0,100	20
	0,050	4,896	5,009	5,109	5,200	5,282	5,358	0,050	
	0,010	5,976	6,092	6,197	6,290	6,375	6,454	0,010	
Error df	Alpha	9	10	11	12	13	14	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/216.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 6 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error 11 y 20 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		15	16	17	18	19	20		
11	0,100	5,310	5,382	5,450	5,514	5,574	5,630	0,100	11
	0,050	5,986	6,063	6,135	6,203	6,266	6,327	0,050	
	0,010	7,567	7,656	7,741	7,820	7,893	7,963	0,010	
12	0,100	5,237	5,308	5,375	5,437	5,495	5,551	0,100	12
	0,050	5,879	5,954	6,024	6,090	6,152	6,210	0,050	
	0,010	7,363	7,448	7,529	7,604	7,674	7,740	0,010	
13	0,100	5,176	5,246	5,311	5,372	5,429	5,483	0,100	13
	0,050	5,790	5,863	5,932	5,996	6,056	6,113	0,050	
	0,010	7,194	7,276	7,354	7,425	7,493	7,557	0,010	
14	0,100	5,124	5,193	5,257	5,317	5,373	5,426	0,100	14
	0,050	5,715	5,786	5,853	5,916	5,974	6,030	0,050	
	0,010	7,052	7,131	7,207	7,276	7,341	7,403	0,010	
15	0,100	5,079	5,147	5,210	5,269	5,324	5,377	0,100	15
	0,050	5,650	5,720	5,786	5,847	5,905	5,959	0,050	
	0,010	6,933	7,008	7,081	7,149	7,212	7,272	0,010	
16	0,100	5,040	5,107	5,169	5,227	5,282	5,334	0,100	16
	0,050	5,594	5,663	5,727	5,787	5,844	5,897	0,050	
	0,010	6,829	6,903	6,974	7,039	7,100	7,159	0,010	
17	0,100	5,006	5,072	5,133	5,191	5,245	5,296	0,100	17
	0,050	5,545	5,612	5,676	5,735	5,791	5,843	0,050	
	0,010	6,739	6,811	6,880	6,944	7,004	7,060	0,010	
18	0,100	4,975	5,040	5,101	5,158	5,212	5,262	0,100	18
	0,050	5,502	5,568	5,630	5,689	5,744	5,796	0,050	
	0,010	6,659	6,730	6,798	6,860	6,918	6,974	0,010	
19	0,100	4,948	5,013	5,073	5,129	5,182	5,232	0,100	19
	0,050	5,463	5,529	5,590	5,648	5,702	5,753	0,050	
	0,010	6,589	6,659	6,725	6,786	6,843	6,898	0,010	
20	0,100	4,924	4,988	5,047	5,103	5,156	5,205	0,100	20
	0,050	5,428	5,493	5,554	5,611	5,664	5,715	0,050	
	0,010	6,527	6,595	6,661	6,720	6,776	6,830	0,010	
Error df	Alpha	15	16	17	18	19	20	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/216.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

**TABLA 7 Valores críticos de la distribución de Rango Studentizado
para grados de libertad de error 30 y 120**

Error df	Alpha	k = number of means or number of steps between ordered means							Alpha	Error df
		2	3	4	5	6	7	8		
30	0,100	2,401	3,018	3,386	3,648	3,852	4,017	4,156	0,100	30
	0,050	2,890	3,488	3,847	4,103	4,302	4,465	4,602	0,050	
	0,010	3,897	4,460	4,804	5,054	5,248	5,406	5,540	0,010	
35	0,100	2,390	3,001	3,365	3,624	3,824	3,986	4,123	0,100	35
	0,050	2,872	3,462	3,815	4,067	4,262	4,422	4,556	0,050	
	0,010	3,859	4,410	4,744	4,987	5,174	5,327	5,457	0,010	
40	0,100	2,382	2,988	3,349	3,605	3,803	3,964	4,099	0,100	40
	0,050	2,859	3,443	3,792	4,040	4,232	4,389	4,521	0,050	
	0,010	3,831	4,372	4,700	4,937	5,120	5,269	5,396	0,010	
45	0,100	2,376	2,979	3,337	3,591	3,787	3,946	4,080	0,100	45
	0,050	2,849	3,429	3,774	4,019	4,210	4,364	4,495	0,050	
	0,010	3,810	4,344	4,666	4,899	5,079	5,225	5,349	0,010	
50	0,100	2,371	2,971	3,327	3,579	3,774	3,932	4,065	0,100	50
	0,050	2,842	3,417	3,759	4,003	4,191	4,344	4,473	0,050	
	0,010	3,793	4,321	4,639	4,868	5,046	5,189	5,312	0,010	
55	0,100	2,367	2,965	3,319	3,570	3,764	3,921	4,053	0,100	55
	0,050	2,835	3,407	3,748	3,990	4,176	4,328	4,456	0,050	
	0,010	3,780	4,302	4,617	4,843	5,019	5,161	5,281	0,010	
60	0,100	2,363	2,959	3,312	3,562	3,755	3,911	4,043	0,100	60
	0,050	2,830	3,399	3,738	3,979	4,164	4,315	4,442	0,050	
	0,010	3,768	4,287	4,599	4,823	4,996	5,137	5,256	0,010	
65	0,100	2,361	2,955	3,307	3,556	3,748	3,903	4,034	0,100	65
	0,050	2,825	3,393	3,730	3,969	4,154	4,303	4,430	0,050	
	0,010	3,759	4,274	4,584	4,805	4,978	5,117	5,235	0,010	
70	0,100	2,358	2,951	3,302	3,550	3,741	3,896	4,026	0,100	70
	0,050	2,822	3,387	3,723	3,961	4,145	4,294	4,419	0,050	
	0,010	3,751	4,263	4,571	4,791	4,962	5,100	5,217	0,010	
120	0,100	2,345	2,931	3,276	3,520	3,708	3,860	3,987	0,100	120
	0,050	2,801	3,357	3,685	3,918	4,097	4,242	4,363	0,050	
	0,010	3,707	4,205	4,501	4,712	4,877	5,010	5,121	0,010	
Error df	Alpha	2	3	4	5	6	7	8	Alpha	Error df
		k = number of means or number of steps between ordered means								

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 7 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error 30 y 120 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		9	10	11	12	13	14		
30	0,100	4,276	4,381	4,475	4,559	4,636	4,706	0,100	30
	0,050	4,721	4,825	4,918	5,001	5,078	5,147	0,050	
	0,010	5,657	5,760	5,853	5,937	6,012	6,082	0,010	
35	0,100	4,241	4,344	4,436	4,519	4,595	4,664	0,100	35
	0,050	4,672	4,774	4,864	4,946	5,020	5,088	0,050	
	0,010	5,570	5,670	5,759	5,840	5,913	5,980	0,010	
40	0,100	4,215	4,317	4,408	4,490	4,564	4,632	0,100	40
	0,050	4,635	4,735	4,825	4,905	4,977	5,044	0,050	
	0,010	5,506	5,603	5,690	5,768	5,840	5,905	0,010	
45	0,100	4,195	4,296	4,386	4,467	4,540	4,607	0,100	45
	0,050	4,607	4,706	4,794	4,873	4,944	5,010	0,050	
	0,010	5,456	5,551	5,636	5,713	5,783	5,847	0,010	
50	0,100	4,179	4,279	4,368	4,448	4,521	4,588	0,100	50
	0,050	4,585	4,682	4,769	4,847	4,918	4,983	0,050	
	0,010	5,417	5,510	5,594	5,669	5,738	5,801	0,010	
55	0,100	4,166	4,265	4,354	4,433	4,506	4,572	0,100	55
	0,050	4,566	4,663	4,749	4,827	4,897	4,961	0,050	
	0,010	5,385	5,477	5,560	5,634	5,701	5,763	0,010	
60	0,100	4,155	4,254	4,342	4,421	4,493	4,558	0,100	60
	0,050	4,551	4,647	4,732	4,809	4,879	4,943	0,050	
	0,010	5,359	5,450	5,531	5,604	5,671	5,732	0,010	
65	0,100	4,146	4,245	4,332	4,410	4,482	4,547	0,100	65
	0,050	4,538	4,634	4,718	4,795	4,864	4,928	0,050	
	0,010	5,337	5,427	5,507	5,580	5,645	5,706	0,010	
70	0,100	4,138	4,236	4,324	4,401	4,472	4,537	0,100	70
	0,050	4,527	4,622	4,706	4,782	4,851	4,914	0,050	
	0,010	5,318	5,408	5,487	5,558	5,624	5,684	0,010	
120	0,100	4,096	4,192	4,277	4,353	4,422	4,485	0,100	120
	0,050	4,468	4,560	4,642	4,715	4,782	4,842	0,050	
	0,010	5,217	5,303	5,378	5,446	5,508	5,565	0,010	
Error df	Alpha	9	10	11	12	13	14	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab..htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

TABLA 7 Valores críticos de la distribución de Rango Studentizado para grados de libertad de error 30 y 120 (Continuación)

Error df	Alpha	k = number of means or number of steps between ordered means						Alpha	Error df
		15	16	17	18	19	20		
30	0,100	4,771	4,831	4,887	4,939	4,988	5,034	0,100	30
	0,050	5,212	5,272	5,328	5,380	5,429	5,476	0,050	
	0,010	6,146	6,206	6,263	6,316	6,366	6,413	0,010	
35	0,100	4,727	4,786	4,841	4,893	4,941	4,986	0,100	35
	0,050	5,151	5,210	5,264	5,315	5,363	5,409	0,050	
	0,010	6,042	6,099	6,154	6,205	6,253	6,298	0,010	
40	0,100	4,695	4,753	4,807	4,857	4,905	4,950	0,100	40
	0,050	5,106	5,163	5,217	5,267	5,314	5,358	0,050	
	0,010	5,965	6,020	6,074	6,123	6,170	6,213	0,010	
45	0,100	4,669	4,727	4,780	4,830	4,877	4,921	0,100	45
	0,050	5,071	5,128	5,180	5,229	5,276	5,319	0,050	
	0,010	5,905	5,960	6,012	6,060	6,105	6,148	0,010	
50	0,100	4,649	4,706	4,759	4,808	4,855	4,899	0,100	50
	0,050	5,043	5,099	5,151	5,200	5,245	5,288	0,050	
	0,010	5,859	5,912	5,964	6,010	6,054	6,096	0,010	
55	0,100	4,633	4,689	4,742	4,791	4,837	4,880	0,100	55
	0,050	5,021	5,076	5,127	5,175	5,220	5,263	0,050	
	0,010	5,821	5,873	5,924	5,969	6,013	6,054	0,010	
60	0,100	4,619	4,675	4,727	4,776	4,822	4,865	0,100	60
	0,050	5,002	5,056	5,107	5,155	5,200	5,242	0,050	
	0,010	5,789	5,841	5,890	5,936	5,979	6,019	0,010	
65	0,100	4,607	4,663	4,715	4,763	4,809	4,852	0,100	65
	0,050	4,986	5,040	5,090	5,138	5,182	5,224	0,050	
	0,010	5,762	5,814	5,862	5,908	5,950	5,990	0,010	
70	0,100	4,597	4,653	4,704	4,752	4,798	4,840	0,100	70
	0,050	4,972	5,026	5,076	5,123	5,167	5,209	0,050	
	0,010	5,739	5,791	5,839	5,884	5,925	5,965	0,010	
120	0,100	4,543	4,597	4,647	4,694	4,738	4,780	0,100	120
	0,050	4,899	4,950	4,999	5,044	5,087	5,127	0,050	
	0,010	5,617	5,666	5,712	5,755	5,794	5,831	0,010	
Error df	Alpha	15	16	17	18	19	20	Alpha	Error df
		k = number of means or number of steps between ordered means							

FUENTE: <http://elvers.stojoe.udayton.edu/psy216/tables/qtab.htm>. Reproducida con la autorización del autor Greg C. Elvers, Ph. D. Para grados de libertad de error superiores, puede consultarse la dirección web aquí señalada.

ANEXO

B

INTRODUCCIÓN DE DATOS EN EL EDITOR DEL SPSS 10.0

- Diseño de dos grupos aleatorios: Capítulo 6, Epígrafe 6.1.2.2.



The screenshot shows the 'dos grupos - Editor de datos SPSS' window. The menu bar includes Archivo, Edición, Ver, Datos, Transformar, Analizar, Gráficos, Utilidades, Ventana, and ?. The toolbar contains icons for file operations, data entry, and analysis. The data grid has 10 rows and 7 columns. The first column is labeled '1: vi' and the second column is labeled '1'. The data is as follows:

	vi	vd	var	var	var	var
1	1,00	5,00				
2	1,00	2,00				
3	1,00	3,00				
4	1,00	1,00				
5	1,00	1,00				
6	2,00	10,00				
7	2,00	13,00				
8	2,00	15,00				
9	2,00	7,00				
10	2,00	8,00				

- **Diseño multigrupos aleatorios: Capítulo 6, Epígrafe 6.2.2.5.**

multigrupos - Editor de datos SPSS						
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?						
1 : tratam 1						
	tratam	recuerdo	var	var	var	var
1	1	15				
2	1	10				
3	1	4				
4	1	8				
5	1	12				
6	2	30				
7	2	32				
8	2	25				
9	2	27				
10	2	35				
11	3	32				
12	3	34				
13	3	40				
14	3	37				
15	3	28				

• **Diseño factorial $A \times B$: Capítulo 7, Epígrafe 7.3.2.4.**

datos factorial Ax8 - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : padres 1

	padres	tipolog	competen	var	var	var
1	1	1	7			
2	1	1	9			
3	1	1	10			
4	1	1	8			
5	1	2	6			
6	1	2	5			
7	1	2	4			
8	1	2	7			
9	2	1	8			
10	2	1	11			
11	2	1	10			
12	2	1	6			
13	2	2	4			
14	2	2	6			
15	2	2	8			
16	2	2	5			
17	3	1	13			
18	3	1	15			
19	3	1	11			
20	3	1	10			
21	3	2	10			
22	3	2	9			
23	3	2	6			
24	3	2	8			

• **Diseño factorial $A \times B \times C$: Capítulo 7, Epígrafe 7.3.3.4.**

factorial Ax8xC - Editor de datos SPSS						
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?						
1 : farmaco						
	farmaco	individu	familiar	estado	var	var
1	1,00	1,00	1,00	22,00		
2	1,00	1,00	1,00	20,00		
3	1,00	1,00	1,00	18,00		
4	1,00	1,00	2,00	9,00		
5	1,00	1,00	2,00	10,00		
6	1,00	1,00	2,00	8,00		
7	1,00	2,00	1,00	14,00		
8	1,00	2,00	1,00	13,00		
9	1,00	2,00	1,00	14,00		
10	1,00	2,00	2,00	35,00		
11	1,00	2,00	2,00	29,00		
12	1,00	2,00	2,00	33,00		
13	2,00	1,00	1,00	16,00		
14	2,00	1,00	1,00	18,00		
15	2,00	1,00	1,00	14,00		
16	2,00	1,00	2,00	34,00		
17	2,00	1,00	2,00	30,00		
18	2,00	1,00	2,00	32,00		
19	2,00	2,00	1,00	28,00		
20	2,00	2,00	1,00	29,00		
21	2,00	2,00	1,00	25,00		
22	2,00	2,00	2,00	12,00		
23	2,00	2,00	2,00	16,00		
24	2,00	2,00	2,00	15,00		

• **Diseño de bloques aleatorios: Capítulo 8, Epígrafe 8.1.2.2.**

bloques aleatorios - Editor de datos SPSS						
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?						
1 : metodol 1						
	metodol	motivaci	rendimie	var	var	var
1	1,00	1,00	4,00			
2	1,00	2,00	7,00			
3	1,00	3,00	8,00			
4	1,00	4,00	12,00			
5	2,00	1,00	14,00			
6	2,00	2,00	12,00			
7	2,00	3,00	15,00			
8	2,00	4,00	19,00			
9	3,00	1,00	2,00			
10	3,00	2,00	4,00			
11	3,00	3,00	5,00			
12	3,00	4,00	5,00			
13	1,00	1,00	3,00			
14	1,00	2,00	6,00			
15	1,00	3,00	3,00			
16	1,00	4,00	11,00			
17	2,00	1,00	12,00			
18	2,00	2,00	10,00			
19	2,00	3,00	14,00			
20	2,00	4,00	13,00			
21	3,00	1,00	1,00			
22	3,00	2,00	2,00			
23	3,00	3,00	4,00			
24	3,00	4,00	3,00			

• **Diseño de cuadrado latino: Capítulo 8, Epígrafe 8.2.2.4.**

cuadrado latino - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1: edad 1

	edad	turno	respons	satisfac	var	var
1	1,00	1,00	1,00	20,00		
2	1,00	2,00	2,00	16,00		
3	1,00	3,00	3,00	12,00		
4	2,00	1,00	2,00	12,00		
5	2,00	2,00	3,00	10,00		
6	2,00	3,00	1,00	14,00		
7	3,00	1,00	3,00	8,00		
8	3,00	2,00	1,00	18,00		
9	3,00	3,00	2,00	10,00		

• **Diseño jerárquico: Capítulo 8, Epígrafe 8.3.2.4.**

jerárquico - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1: programa 1

	programa	centro	conducta	var	var	var
1	1,00	1,00	12,00			
2	1,00	1,00	12,00			
3	1,00	1,00	13,00			
4	1,00	2,00	17,00			
5	1,00	2,00	13,00			
6	1,00	2,00	20,00			
7	2,00	3,00	11,00			
8	2,00	3,00	15,00			
9	2,00	3,00	12,00			
10	2,00	4,00	5,00			
11	2,00	4,00	9,00			
12	2,00	4,00	12,00			
13	3,00	5,00	4,00			
14	3,00	5,00	1,00			
15	3,00	5,00	6,00			
16	3,00	6,00	12,00			
17	3,00	6,00	8,00			
18	3,00	6,00	10,00			

- **Diseño con covariables: Capítulo 8, Epígrafe 8.4.3.4.**

covarianza - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : pretest 7

	pretest	posttest	hora	var	var	var
1	7,00	9,00	1,00			
2	6,00	10,00	1,00			
3	4,00	7,00	1,00			
4	8,00	9,00	1,00			
5	5,00	8,00	1,00			
6	8,00	12,00	2,00			
7	7,00	8,00	2,00			
8	6,00	7,00	2,00			
9	9,00	11,00	2,00			
10	7,00	10,00	2,00			

- **Diseño intrasujeto simple (Diseño $A \times S$): Capítulo 9, Epígrafe 9.1.2.2.**

AxS - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : a1 7

	a1	a2	a3	var	var	var
1	7,00	4,00	5,00			
2	8,00	5,00	7,00			
3	8,00	5,00	3,00			
4	9,00	4,00	2,00			
5	10,00	3,00	6,00			
6	7,00	6,00	5,00			

- **Diseño intrasujeto de dos factores (Diseño $A \times B \times S$): Capítulo 9, Epígrafe 9.1.2.3.**

AxBxS - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1: casprin 19

	ingprin	ingsecun	ingejem	eusprin	eussecun	eusejem
1	19,00	25,00	26,00	19,00	14,00	8,00
2	26,00	16,00	31,00	16,00	11,00	9,00
3	25,00	10,00	17,00	13,00	20,00	13,00

- **Diseños de medidas parcialmente repetidas: Diseño *split-plot*: Capítulo 9, Epígrafe 9.2.2.6.**

split-plot - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1:

	sexo	sust	adj	var	var	var
1	,00	9,00	11,00			
2	,00	4,00	4,00			
3	,00	6,00	10,00			
4	1,00	12,00	7,00			
5	1,00	14,00	12,00			
6	1,00	9,00	14,00			

- **Diseño *cross-over* o conmutativo:** Capítulo 9, Epígrafe 9.3.1.2.

cross-over - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : tratam 1

	tratam	per1	per2	var	var	var
1	1,00	5,00	4,00			
2	1,00	6,00	4,00			
3	1,00	4,00	2,00			
4	1,00	5,00	3,00			
5	1,00	6,00	5,00			
6	2,00	4,00	7,00			
7	2,00	3,00	5,00			
8	2,00	3,00	4,00			
9	2,00	5,00	6,00			
10	2,00	2,00	4,00			

- **Diseño de cuadrado latino intrasujeto:** Capítulo 9, Epígrafe 9.3.2.2.

cuadrado latino intrasujeto - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : fam 1

	fam	sujeto	orden	recuerdo	var	var
1	1,00	1,00	1,00	7,00		
2	2,00	1,00	2,00	5,00		
3	3,00	1,00	3,00	4,00		
4	2,00	2,00	1,00	5,00		
5	3,00	2,00	2,00	4,00		
6	1,00	2,00	3,00	6,00		
7	3,00	3,00	1,00	3,00		
8	1,00	3,00	2,00	6,00		
9	2,00	3,00	3,00	4,00		

- **Diseño experimental multivariado: Capítulo 10, Epígrafe 10.1.2.3.**

multivariante - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : edad 1

	edad	metodos	satisfac	var	var	var
1	1,00	8,00	8,00			
2	1,00	6,00	7,00			
3	1,00	7,00	6,00			
4	2,00	11,00	5,00			
5	2,00	7,00	6,00			
6	2,00	9,00	4,00			
7	3,00	6,00	3,00			
8	3,00	5,00	2,00			
9	3,00	4,00	4,00			

- **Diseño de bloques incompletos: Capítulo 10, Epígrafe 10.2.3.3.**

bloques incompletos - Editor de datos SPSS

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?

1 : bloques 1

	bloques	sexo	barrio	replica	fobia	var
1	1	1	1	1	5	
2	1	2	2	1	4	
3	1	1	1	2	4	
4	1	2	2	2	5	
5	1	1	1	3	6	
6	1	2	2	3	4	
7	2	2	1	1	2	
8	2	1	2	1	7	
9	2	2	1	2	4	
10	2	1	2	2	5	
11	2	2	1	3	3	
12	2	1	2	3	6	

- **Diseño factorial fraccionado: Capítulo 10, Epígrafe 10.3.3.3.**

factorial fraccionado - Editor de datos SPSS						
Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana ?						
1: a 2						
	a	b	c	vd	var	var
1	2,00	1,00	1,00	8,00		
2	2,00	1,00	1,00	6,00		
3	2,00	1,00	1,00	5,00		
4	1,00	2,00	1,00	3,00		
5	1,00	2,00	1,00	2,00		
6	1,00	2,00	1,00	3,00		
7	1,00	1,00	2,00	9,00		
8	1,00	1,00	2,00	8,00		
9	1,00	1,00	2,00	10,00		
10	2,00	2,00	2,00	5,00		
11	2,00	2,00	2,00	6,00		
12	2,00	2,00	2,00	4,00		

REFERENCIAS BIBLIOGRÁFICAS

- Ander-Egg, E. (1990). *Técnicas de investigación social*. Buenos Aires: Humanitas.
- Anguera, M. T. (1981). La observación (I): Problemas metodológicos. En R. Fernández Ballesteros y J. A. I. Carrobbles (Eds.), *Evaluación conductual: Metodología y aplicaciones* (pp. 292-333). Madrid: Pirámide.
- Anguera, M. T. (1983). *Manual de prácticas de observación*. México: Trillas.
- Anguera, M. T. (1990). Metodología observacional. En J. Arnau, M. T. Anguera y J. Gómez (Eds.), *Metodología de la investigación en ciencias del comportamiento* (pp. 123-236). Murcia: Universidad de Murcia.
- Anguera, M. T. (1991a). Análisis del experimento desde la metodología científica. En J. Pascual, M. T. Anguera, G. Vallejo y F. Salvador (Eds.), *Psicología experimental* (pp. 107-155). Valencia: Nau Llibres.
- Anguera, M. T. (Ed.) (1991b). *Metodología observacional en la investigación psicológica. Vol. I: Fundamentación (1)*. Barcelona: P.P.U.
- Arnau, J. (1978a). *Psicología experimental: Un enfoque metodológico*. México: Trillas.
- Arnau, J. (Ed.) (1978b). *Métodos de investigación en las ciencias humanas*. Barcelona: Omega.
- Arnau, J. (1986). *Diseños experimentales en psicología y educación (vol. I)*. México: Trillas.
- Arnau, J. (1989). Metodología de la investigación y diseños. En J. Mayor y J. L. Pinillos (Eds.), *Tratado de psicología general. Vol. I: Teoría, historia y método* (Coords. J. Arnau y H. Carpintero) (pp. 581-616). Madrid: Alhambra.
- Arnau, J. (1990a). *Diseños experimentales multivariados*. Madrid: Alianza.
- Arnau, J. (1990b). Metodología experimental. En J. Arnau, M. T. Anguera y J. Gómez (Eds.), *Metodología de la investigación en ciencias del comportamiento* (pp. 7-122). Murcia: Universidad de Murcia.
- Arnau, J. (1994). *Diseños experimentales en esquemas*. Vol. I. Barcelona: Universidad de Barcelona.
- Arnau, J. (1995b). Fundamentos metodológicos de los diseños experimentales de sujeto único. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Eds.), *Métodos de investigación en psicología* (pp. 163-177). Madrid: Síntesis.

- Arnau, J. (1995e). *Diseños longitudinales aplicados a las ciencias sociales y del comportamiento*. México: Limusa.
- Arnau, J., Anguera, M. T. y Gómez, J. (1990). *Metodología de la investigación en ciencias del comportamiento*. Murcia: Universidad de Murcia.
- Ato, M. (1991). *Investigación en ciencias del comportamiento I: Fundamentos*. Barcelona: P.P.U.
- Ato, M. y López, J. J. (1992). Análisis de covarianza en diseños de medidas repetidas: El riesgo de una interpretación. *Anuario de Psicología*, 55, 91-108.
- Ato, M., López, J. J. y Serrano, J. M. (1981). *Fundamentos de estadística inferencial*. Murcia: Librería Yerba.
- Balluerka, N. (1997). *Proyecto docente. Memoria de oposición a Titularidad de Universidad*. Universidad del País Vasco/Euskal Herriko Unibertsitatea.
- Balluerka, N. (1999). *Planificación de la investigación: La validez del diseño*. Salamanca: Ediciones Amarú.
- Balluerka, N. e Isasi, X. (1996). *Saiakuntza eta sasisaiakuntza diseinuaren balidezia psikologian*. Bilbao: U.E.U.
- Barcikowsky, R. S. (1981). Statistical power with group mean as the unit of analysis. *Journal of Educational Statistics*, 6, 267-285.
- Bock, R. D. (1975). *Multivariate statistical methods in behavioral research*. New York, NY: McGraw-Hill.
- Bock, R. D. y Haggard, E. A. (1968). The use of multivariate analysis of variance in behavioral research. En D. Whitla (Ed.), *Handbook of measurement and assesment sciences*. Reading, MA: Addison-Wesley.
- Box, G. E. P. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems: II. Effect on inequality of variance and correlation of errors in the two-way classification. *Annals of Mathematical Statistics*, 25, 484-498.
- Bray, J. H. y Maxwell, S. E. (1982). Analyzing and interpreting significant MANOVAs. *Review of Educational Research*, 52, 340-367.
- Bray, J. H. y Maxwell, S. E. (1993). Multivariate analysis of variance. En M.S. Lewis-Beck (Ed.), *Experimental design and methods* (pp. 337-404). London: Sage.
- Brown, M. B. y Forsythe, A. B. (1974). The ANOVA and multiple comparisons for data with heterogeneous variances. *Biometrics*, 30, 719-724.
- Brown, S. R. y Melamed, L. E. (1993). Experimental design and analysis. En M. S. Lewis-Beck (Ed.), *Experimental design and methods* (pp. 75-158). London: Sage.
- Bryan, J. L. y Paulson, A. S. (1976). An extension of Tukey's method of multiple comparisons for data with random concomitant variables. *Biometrika*, 36, 69-79.
- Campbell, D. T. y Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. En N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 171-246). Chicago: Rand McNally.
- Campbell, D. T. y Stanley, J. C. (1988). *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu (Ed. orig. 1966).
- Chambers, J. M., Cleveland, W. S., Kleiner, B. y Tukey, P. A. (1983). *Graphical methods for data analysis*. Belmont, CA: Wadsworth.
- Christensen, L. B. (1988). *Experimental methodology* (4th ed.). Boston, MA: Allyn and Bacon.
- Cicchetti, D. V. (1972). Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin*, 77, 405-408.
- Clinch, J. J. y Keselman, H. J. (1982). Parametric alternatives to the analysis of variance. *Journal of Educational Statistics*, 7, 207-214.
- Cochran, W. G. y Cox, G. M. (1957). *Experimental designs*. New York, NY: Wiley.
- Cohen, J. y Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York, NY: Wiley.
- Cook, T. D. (1983). Quasiexperimentation. En G. Mongan (Ed.), *Beyond methods*. London: Sage.
- Cook, T. D. y Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand McNally.

- Cook, T. D. y Campbell, D. T. (1986). The causal assumptions of quasiexperimental practice. *Synthese*, 68, 141-160.
- Cook, T. D., Campbell, D. T. y Peracchio, L. (1990). Quasi-experimentation. En M. D. Dunnette y L. M. Houghs (Eds.), *Handbook of industrial and organizational psychology* (pp. 491-576). Palo Alto, CA: Consulting Psychologist Press.
- Cook, T. D. y Reichardt, C. S. (1982). *Qualitative and quantitative methods in evaluation research*. Beverly Hills, CA: Sage.
- Cooley, W. W. y Lohnes, P. R. (1971). *Multivariate data analysis*. New York, NY: John Wiley.
- Cox, D. R. (1957). The use of a concomitant variable in selecting an experimental design. *Biometrika*, 44, 150-158.
- Cozby, P. C. (1993). *Methods in behavioral research* (5th ed.). Mountain View, CA: Mayfield.
- Cramer, E. M. y Bock, R. D. (1966). Multivariate analysis. *Review of Educational Research*, 36, 604-617.
- Cronbach, L.J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671-684.
- Darlington, R. B., Weinberg, S. L. y Walberg, H. J. (1973). Canonical variate analysis and related techniques. *Review of Educational Research*, 43, 443-454.
- Davidson, M. L. (1972). Univariate versus multivariate tests in repeated measures experiments. *Psychological Bulletin*, 77, 446-452.
- Domenech, J. M. y Riba, M. D. (1985). *Métodos estadísticos: Modelo lineal de regresión*. Barcelona: Herder.
- Dowdy, S. y Wearden, S. (1991). *Statistics for research (2nd ed.) (Wiley series in probability and mathematical statistics)*. New York, NY: Wiley.
- Draper, N. y Smith, H. (1981). *Applied Regression Analysis*. 2nd. Edition. New York, NY: John Wiley and Sons.
- Duncan, D. B. (1955). Multiple range and multiple F tests. *Biometrics*, 11, 1-42.
- Dunham, P. J. (1988). *Research methods in psychology*. New York, NY: Harper and Row.
- Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56, 52-64.
- Dunn, O. J. y Clark, V. A. (1974). *Applied statistics: Analysis of variance and regression*. New York, NY: Wiley.
- Dunnett, C. W. (1955). A multiple comparisons procedure for comparing several treatments with a control. *Journal of the American Statistical Association*, 50, 1096-1121.
- Dwyer, J. H. (1983). *Statistical models for the social and behavioral sciences*. New York, NY: Oxford University Press.
- Edwards, A. L. (1985). *Experimental design in psychological research* (5th ed.). New York, NY: Holt, Rinehart and Winston.
- Elashoff, J. D. (1969). Analysis of covariance: A delicate instrument. *American Educational Research Journal*, 6, 383-401.
- Emerson, J. D. y Stoto, M. A. (1983). Transforming data. En D. C. Floagly, F. Mosteller y J. W. Tukey (Eds.), *Understanding robust and exploratory data analysis* (pp. 97-128). New York, NY: John Wiley.
- Federer, W. T. (1955). *Experimental design. Theory and application*. New York, NY: Macmillan.
- Feldt, L. S. (1958). A comparison of the precision of three experimental design employing a concomitant variable. *Psychometrika*, 23, 335-354.
- Festinger, L. y Katz, D. (Eds.) (1953). *Research methods in the behavioral sciences*. New York, NY: Dryden Press.
- Finn, J. D. (1974). *A general model for multivariate analysis*. New York, NY: Holt, Rinehart and Winston.
- Fisher, R. A. (1925). *Statistical methods for research workers*. London: Oliver and Boyd.
- Fisher, R. A. (1935). *The design of experiments*. London: Oliver and Boyd.
- Fisher, R. A. y Yates, F. (1953). *Statistical tables for biological, agricultural and medical research* (4th ed.). Edimburgo: Oliver and Boyd.

- García, M. V. (1992). *El método experimental en la investigación psicológica*. Barcelona: P.P.U.
- Girden, E. R. (1992). *Anova: Repeated measures*. London: Sage.
- Glass, G. V., Peckham, P. D. y Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed-effects analysis of variance and covariance. *Review of Educational Research*, 42, 237-288.
- Gómez, J. (1990). Metodología de encuesta por muestreo. En J. Arnau, M. T. Anguera y J. Gómez (Eds.), *Metodología de la investigación en ciencias del comportamiento* (pp. 237-310). Murcia: Universidad de Murcia.
- Green, P. E. (1978). *Analyzing multivariate data*. Hillsdale, IL: Dryden Press.
- Green, P. E. y Carroll, J. D. (1976). *Mathematical tools for applied multivariate analysis*. New York, NY: Academic Press.
- Greenhouse, S. W. y Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, 24, 95-112.
- Grupo ModEst (2000a). *Análisis de datos: Del contraste de hipótesis al modelado estadístico*. Barcelona: C.B.S.
- Grupo ModEst (2000b). *El Modelo Lineal Generalizado*. Barcelona: C.B.S.
- Haimson, B. R. y Elfenbeim, M. H. (1985). *Experimental methods in psychology*. New York, NY: McGraw-Hill.
- Harris, R. J. (1975). *A primer of multivariate statistics*. New York, NY: Academic Press.
- Hochberg, Y. y Tamhane, A. C. (1987). *Multiple comparison procedures*. New York, NY: John Wiley and Sons.
- Howell, D. C. (1987). *Statistical methods for psychology* (2nd ed.). Boston: PWS-Kent.
- Hsu, J. C. (1996). *Multiple comparisons. Theory and methods*. London: Chapman & Hall.
- Huberty, C. J. (1975). Discriminant analysis. *Review of Educational Research*, 45, 543-598.
- Huberty, C. J. (1984). Issues in the use and interpretation of discriminant analysis. *Psychological Bulletin*, 95, 156-171.
- Huitema, B. E. (1980). *The analysis of covariance and alternatives*. New York, NY: John Wiley and Sons.
- Hummel, T. J. y Sligo, J. R. (1971). Empirical comparison of univariate and multivariate analyses of variance procedures. *Psychological Bulletin*, 75, 49-57.
- Huynh, H. y Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurement designs have exact F-distributions. *Journal of the American Statistical Association*, 63, 1582-1589.
- Jaccard, J., Becker, M. A. y Wood, G. (1984). Pairwise multiple comparison procedures: A review. *Psychological Bulletin*, 93, 589-596.
- Jones, B. y Kenward, M. G. (1989). *Design and analysis of cross-over trials*. New York, NY: Chapman and Hall.
- Judd, C. M. y McClelland, G. H. (1989). *Data analysis: a model comparison approach*. San Diego, CA: Harcourt, Brace and Jovanovich.
- Kantowitz, B. H., Roediger, H. L. y Elmes, D. G. (1991). *Experimental psychology. Understanding psychological research* (4th ed.). St. Paul, MN: West.
- Kazdin, A. E. (1992). *Research designs in clinical psychology*. Massachusetts: Allyn & Bacon.
- Kenny, D. A. (1979). *Correlation and causality*. New York, NY: Wiley.
- Kenny, D. A. y Judd, C. A. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, 99, 422-431.
- Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall Inc.
- Keppel, G. y Zedeck, S. (1989). *Data analysis for research. Analysis of variance and multiple regression/correlation approaches*. New York, NY: Freeman and Company.
- Kerlinger, F. N. (1979). *Behavioral research: A conceptual approach*. New York, NY: Holt, Rinehart and Winston.
- Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd ed.). New York, NY: Holt, Rinehart and Winston.

- Kerlinger, F. N. y Pedhazur, E. J. (1973). *Multiple regression in behavioral research*. New York, NY: Holt, Rinehart and Winston.
- Keselman, H. J. (1982). Multiple comparisons for repeated measures means. *Multivariate Behavioral Research*, 17, 87-92.
- Keselman, H. J., Rogan, C. J., Mendoza, J. L. y Breen, L. J. (1980). Testing the validity conditions of repeated measures F tests. *Psychological Bulletin*, 87, 479-481.
- Keuls, M. (1952). The use of studentized range in connection with an analysis of variance. *Euphytica*, 1, 112-122.
- Kiess, H. O. y Bloomquist, D. W. (1985). *Psychological research methods. A conceptual approach*. Boston, MA: Allyn and Bacon.
- Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Belmont, CA: Brooks/Cole.
- Kish, L. (1987). *Statistical design for research*. New York, NY: Wiley.
- Klockars, A. J. y Sax, A. (1985). *Multiple comparisons. Series: Quantitative applications in the social sciences*. Vol. 61. Beverly Hills, CA: Sage Publications.
- Klockars, A. J. y Sax, G. (1986). *Multiple comparisons*. Beverly Hills: Sage.
- Levin, J. R. y Marascuilo, L. A. (1972). Type IV error and interactions. *Psychological Bulletin*, 78, 368-374.
- Lewis-Beck, M.S. (1980). Applied regression: An introduction. *Sage University Paper series on Quantitative Applications in the Social Sciences*, 07-022. Beverly Hills, CA: Sage.
- Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov Test for normality with mean and variance unknown. *Journal of American Statistical Association*, 64, 399-402.
- López, J. J. (1995). *Proyecto docente. Memoria de oposición a Titularidad de Universidad*. Universidad de Murcia.
- Macnaughton, D. B. (1998). Which Sums of Squares are the best in unbalanced analysis of variance? Artículo publicado en la dirección URL <http://matstat.com/ss/>.
- Marascuilo, L. A. y Busk, P. L. (1988). Combining statistics for multiple-baseline AB and replicated ABAB designs across subjects. *Behavioral Assessment*, 10, 1-28.
- Martínez Arias, M. R. (1983). Métodos de investigación en psicología evolutiva. En A. Marchesi, M. Carretero y J. Palacios (Eds.), *Psicología evolutiva. 1. Teoría y métodos* (pp. 354-368). Madrid: Alianza.
- Martínez Arias, M. R. (1986). Métodos de investigación en la psicología ambiental. En F. Jiménez Burillo y J. I. Aragonés (Comp.), *Introducción a la psicología ambiental* (pp. 331-366). Madrid: Alianza.
- Mauchly, J. W. (1940). Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11, 204-209.
- Maxwell, S. E. (1980). Pairwise multiple comparisons in repeated measures designs. *Journal of Educational Statistics*, 5, 269-287.
- Maxwell, S. E. y Delaney, H. D. (1990). *Designing experiments and analyzing data. A model comparison perspective*. California: Brooks/Cole Publishing Company.
- McCall, R. B. y Appelbaum, M. I. (1973). Bias in the analysis of repeated measures designs: Some alternative approaches. *Child Development*, 44, 401-415.
- Mead, R. (1988). *The design of experiments: Statistical principles for practical applications*. Cambridge, MA: Cambridge University Press.
- Mendoza, J. L. (1980). A significance test for multisample sphericity. *Psychometrika*, 45, 495-498.
- Meredith, W. (1964). Canonical correlation with fallible data. *Psychometrika*, 29, 55-65.
- Meyer, D. L. (1991). Misinterpretation of interaction effects: A reply of Rosnow and Rosenthal. *Psychological Bulletin*, 110, 571-573.
- Meyers, L. S. y Grossen, N. E. (1974). *Behavioral research: Theory, procedure and design*. San Francisco, CA: Freeman.
- Milliken, G. A. y Johnson, D. E. (1984). *Analysis of Messy Data. Vol I: Designed experiments*. New York, NY: Van Nostrand Reinhold Company.
- Mitzel, C. H. y Games, P. A. (1981). Circularity and multiple comparisons in repeated measures designs. *British Journal of Mathematical and Statistical Psychology*, 34, 253-259.

- Moreno, R. y López, J. (1985). *Análisis metodológico de investigaciones experimentales en psicología*. Barcelona: Alamex.
- Morran, D. K., Robinson, F. F. y Hulse, K. D. (1990). Group research and the unit of analysis problem: The use of ANOVA designs with nested factors. *Journal of Specialists in Group Work*, 15, 10-14.
- Morrison, D. F. (1967). *Multivariate statistical methods*. New York, NY: McGraw-Hill.
- Namboodiri, W. K. (1984). *Matrix Algebra: An Introduction*. Beverly Hills, CA: Sage.
- Namboodiri, W. K., Carter, L. F. y Blalock, H. M., Jr. (1975). *Applied multivariate analysis and experimental designs*. New York, NY: McGraw-Hill.
- Newman, D. (1939). The distribution of range in samples of a normal population, expressed in terms of independent estimate of a standard deviation. *Biometrika*, 31, 20-30.
- O'Neil, W. M. (1962). *An introduction to method in psychology* (2nd ed.). Australia: Melbourne University Press.
- Olson, C. L. (1974). Comparative robustness of six tests in multivariate analysis of variance. *Journal of the American Statistical Association*, 69, 894-908.
- Olson, C.L. (1976). On Choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 83, 579-586.
- Ottensbacher, K. J. (1991). Interpretation of interaction in factorial analysis of variance design. *Statistics in Medicine*, 10, 1465-1571.
- Pardo, A. y San Martín, R. (1994). *Análisis de datos en psicología II*. Madrid: Pirámide.
- Pascual, J. (1995a). Diseño entre grupos. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Eds.), *Métodos de investigación en psicología* (pp. 73-107). Madrid: Síntesis.
- Pascual, J. (1995b). Diseño de covarianza, bloqueo y doble bloqueo. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual y G. Vallejo (Eds.), *Métodos de investigación en psicología* (pp. 137-158). Madrid: Síntesis.
- Pascual, J. (1995c). Diseño de medidas repetidas. En M. T. Anguera, J. Arnau, M. Ato, R. Martínez, J. Pascual, y G. Vallejo (Eds.), *Métodos de investigación en psicología* (pp. 113-134). Madrid: Síntesis.
- Pascual, J. y Camarasa, C. (1991). Diseños: Supuestos, potencia y tamaño del efecto. En J. Pascual, M. T. Anguera, G. Vallejo y F. Salvador (Eds.), *Psicología experimental* (pp. 75-106). Valencia: Nau Llibres.
- Pascual, J., Frías, D. y García, F. (1996). *Manual de psicología experimental*. Barcelona: Ariel.
- Pascual, J., García, J. F. y Frías, M. D. (1995). *El diseño y la investigación experimental en psicología*. Valencia: C.S.V.
- Pascual, J. y Musitu, G. (1981). Crisis y sentido de la experimentación de laboratorio. *Psicológica*, 2 (2), 143-154.
- Pedhazur, E. J. y Pedhazur Schmelkin, L. (1991). *Measurement, design and analysis. An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Peña, D. (1987). *Estadística, modelos y métodos. Vol. II: Modelos lineales y series temporales*. Madrid: Alianza.
- Pereda, S. (1987). *Psicología experimental I: Metodología*. Madrid: Pirámide.
- Porebski, O. R. (1966a). On the interrelated nature of the multivariate statistics used in discriminatory analysis. *British Journal of Mathematical and Statistical Psychology*, 19, 197-214.
- Porebski, O. R. (1966b). Discriminatory and canonical analysis of technical college data. *British Journal of Mathematical and Statistical Psychology*, 19, 215-236.
- Postman, L. (1955). The probability approach and nomothetic theory. *Psychological Review*, 62, 218-255.
- Putnam, H. (1981). The corroboration of theories. En I. Hacking (Ed.), *Scientific revolutions*. Oxford: Oxford University Press.
- Ramsey, P. H. (1981). Power of univariate pairwise multiple comparison procedures. *Psychological Bulletin*, 90 (2), 352-366.

- Rao, C. R. (1951). An asymptotic expansion of the distribution of Wilks' criterion. *Bulletin of the International Statistics Institute*, 33, 177-180.
- Reaves, C. C. (1992). *Quantitative research for the behavioral sciences*. New York, NY: Wiley.
- Riba, M. D. (1990). *Modelo lineal de análisis de la variancia*. Barcelona: Herder.
- Robinson, P. W. (1976). *Fundamentals of experimental psychology*. Englewood Cliffs, NJ: Prentice-Hall.
- Rogan, J. C. y Keselman, H. J. (1977). Is the ANOVA F-test robust to variance heterogeneity when sample sizes are equal? An investigation via a coefficient of variation. *American Educational Research Journal*, 14, 493-498.
- Rogosa, D. R. (1980). Comparing non-parallel regression lines. *Psychological Bulletin*, 88, 307-321.
- Rosenthal, R. y Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York, NY: McGraw-Hill.
- Rosnow, R. L. y Rosenthal, R. (1989). Definition and interpretation of interaction effects. *Psychological Bulletin*, 105, 143-146.
- Rouanet, H. y Lepine, D. (1970). Comparison between treatments in a repeated measurement design: ANOVA and multivariate methods. *British Journal of Mathematical and Statistical Psychology*, 23, 147-163.
- Roy, J. (1958). Step-down procedure in multivariate analysis. *Annals of Mathematical Statistics*, 29, 1177-1187.
- Scariano, S. M. y Davenport, J. M. (1987). The effects of violation of independence assumptions in the one-way ANOVA. *The American Statistician*, 41, 123-129.
- Scheffé, H. (1959). *The analysis of variance*. New York, NY: John Wiley and Sons.
- Seaman, M. A., Levin, J. R. y Serlin, R. C. (1991). New developments in pairwise multiple comparisons: some powerful and practicable procedures. *Psychological Bulletin*, 3, 577-586.
- Searle, S. R. (1987). *Linear models for unbalanced data*. New York: John Wiley.
- Shaughnessy, J. J. y Zechmeister, E. B. (1994). *Research methods in psychology* (3rd ed.). Singapore: McGraw-Hill.
- Siegel, S. (1956). *Non parametric statistics for the behavioral sciences*. New York, NY: McGraw-Hill.
- Spector, P. E. (1977). What to do with significant multivariate effects in multivariate analyses of variance. *Journal of Applied Psychology*, 62, 158-163.
- Spector, P. E. (1993). Research designs. En M.S. Lewis-Beck (Ed.), *Experimental design and methods* (pp. 1-72). London: Sage.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Stevens, J. P. (1972). Four methods of analyzing between variation for the K-group MANOVA problem. *Multivariate Behavioral Research*, 7, 499- 522.
- Stevens, J. P. (1973). Step-down analyses and simultaneous confidence intervals in MANOVA. *Multivariate Behavioral Research*, 8, 391-402.
- Strahan, R. I. (1982). Multivariate analysis and the problem of Type I error. *Journal of Counseling Psychology*, 29, 175-179.
- Suppes, P. (1970). *A probabilistic theory of causality*. Amsterdam: North Holland.
- Suter, W. N., Lindgren, H. C. y Hiebert, S. J. (1989). *Experimentation in psychology: A guided tour*. Boston, MA: Allyn & Bacon.
- Tabachnik, B. G. y Fidell, L. S. (1989). *Using multivariate statistics* (2nd. ed.). New York, NY: Harper and Row.
- Tatsuoka, M. M. (1971). *Multivariate analysis: Techniques for educational and psychological research*. New York, NY: John Wiley.
- Timm, N. H. (1975). *Multivariate analysis with applications in education and psychology*. Monterey, CA: Brooks/Cole.
- Tomarken, A. J. y Serlin, R. C. (1986). Comparison of ANOVA alternatives under variance heterogeneity and specific noncentrality structures. *Psychological Bulletin*, 99, 90-99.
- Toothaker, L. E. (1991). *Multiple comparisons for researchers*. Newbury Park, CA: Sage.
- Tukey, J. W. (1949). One degree of freedom for nonadditivity. *Biometrics*, 5, 232-242.

- Tukey, J. W. (1953). *The problem of multiple comparisons*. Unpublished manuscript, Princeton University.
- Ury, H. K. (1976). A comparison of four procedures for multiple comparisons among means (pairwise contrasts) for arbitrary sample sizes. *Technometrics*, 18, 89-97.
- Vallejo, G. (1986a). Procedimientos simplificados de análisis en los diseños de series temporales interrumpidas: Modelos estáticos. *Revista de Terapia del Comportamiento*, 4 (2), 114-148.
- Vallejo, G. (1986b). Procedimientos simplificados de análisis en los diseños de series temporales interrumpidas: Modelos dinámicos. *Revista de Terapia del Comportamiento*, 4 (3), 323-349.
- Vallejo, G. (1991). *Análisis univariado y multivariado de los diseños de medidas repetidas de una sola muestra y de muestras divididas*. Barcelona: P.P.U.
- Vinod, H. D. y Ullah, A. (1981). *Recent advances in regression methods*. New York, NY: Marcel Dekker, Inc.
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38, 330-336.
- Wilcox, R. R. (1987). New designs in analysis of variance. *Annual Review of Psychology*, 38, 29-60.
- Wilcox, R. R., Charlin, V. L. y Thompson, K. L. (1986). New Monte Carlo results on the robustness of the ANOVA F, W and F* statistics. *Communications in Statistics-Simulation and Computation*, 15, 933-943.
- Wilkinson, L. (1975). Response variable hypothesis in the multivariate analysis of variance. *Psychological Bulletin*, 82, 408-412.
- Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York, NY: McGraw-Hill.
- Winer, B. J., Brown, D. R. y Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York, NY: McGraw-Hill.
- Zimny, G. H. (1961). *Method in experimental psychology*. Chicago, IL: Ronald Press.

ÍNDICE DE TÉRMINOS

A

Aleatorización (*véase* regla de asignación aleatoria)

Alpha

nominal, 46, 246

real, 46, 246

Análisis

de la covarianza, 24, 41, 157, 208-239
 modelo de la regresión, 228-234
 supuestos básicos del, 210-213, 216-222, 235-237

de la regresión, 209, 228

de la varianza, 27-34, 44

 aplicado a la regresión, 34

 factorial, 92, 116, 159, 164, 181, 191

 mixto, 243-278

 para los diseños de medidas

 parcialmente repetidas, 283-302

 supuestos básicos, 283

 supuestos básicos del, 243-246

 multivariado, 34-42, 322-332

 unidireccional de Kruskal-Wallis, 58

 unifactorial, 48, 58

 univariado, 27-34

 ecuación estructural, 28, 50

 modelo estructural matemático, 27

discriminante, 37, 41

step-down, 41

ANCOVA (*véase* análisis de la covarianza)

Anidamiento

técnica de, 23, 157, 190

ANOVA (*véase* análisis de la varianza)

Asignación

aleatoria (*véase* regla de asignación aleatoria)

Autovalores, 38, 326, 328

Autovector, 38

B

Bartlett

prueba de, 46

Bloqueo

técnica de, 23, 157-158, 180, 242

Bonferroni (*véase también* comparaciones múltiples)

procedimiento o corrección de, 34, 41, 68, 70, 71, 77-81, 84, 150, 152-153, 280

C

C de Dunnett (*véase* Dunnett)

Carry-over (*véase* efectos *carry-over*)

Cicchetti

corrección de, 154-155

Cochran

prueba de, 46

Coefficiente de determinación, 209

Combinaciones experimentales

configuración completa de las, 89

configuración incompleta de las, 89

Comparaciones múltiples, 34, 42, 67-88, 94, 111-112, 151-156, 279-282, 296-298

a posteriori (*véase* comparaciones múltiples no planificadas)

a priori (*véase* comparaciones múltiples planificadas)
 complejas, 68, 84, 281
 entre pares de medias (*véase* comparaciones simples)
 no planificadas, 34, 67, 68, 83, 234, 279, 281
 planificadas, 34, 67, 81-82, 84, 151, 279, 280, 281
 simples, 68, 81-84, 83, 279, 281

Contrabalanceo, 242, 313

Contrastes

de hipótesis específicas univariadas, 41
 de medias (*véase* comparaciones múltiples)
 no ortogonales (*véase también* comparaciones múltiples planificadas), 34, 42, 67
 ortogonales (*véase también* comparaciones múltiples planificadas), 34, 42, 67, 80

Control, 6, 8-9, 13-16, 20

estadístico, 21, 157, 209

Corrección de Bonferroni (*véase* Bonferroni)

Correlación canónica, 38

Correlacional

estudio, 7
 Psicología, 7

Covariable, 208

Criterio

de la raíz más grande, 39
 de la razón de verosimilitud, 38
 de la traza, 39

Cuasi-experimento, 14

D

Desviación

explicada, 30
 no explicada, 30

Diseño(s)

alternativos (*véase* diseños *cross-over*)
 bivalente, 22, 25
 complejos, 22, 25
 completos, 23, 26
 con grupo
 de control, 43
 acoplado, 44
 sin contacto, 44
 de lista de espera, 44
 placebo, 43
 con covariables (*véase también* diseños de covarianza), 208-239
 con variables anidadas (*véase* diseño jerárquico)
 conmutativos (*véase* diseños *cross-over*)

cross-over, 242, 303-313
 cuasi-experimental, 7-10, 13, 16
 de bloques, 23-25
 aleatorios, 23, 25, 158-179
 con más de un sujeto por casilla, 159, 164-179
 con un solo sujeto por casilla, 159-164
 incompletos, 24, 26, 42, 332-346
 de comparación de grupos, 21
 de control, 282
 de covarianza, 24, 26, 42
 de cuadrado
 grecolatino, 23, 25, 26, 180
 hipergrecolatino, 180
 latino, 23, 25
 intersujetos, 23, 26, 180-190
 intrasujeto, 242, 313-320
 de dos grupos, 22, 25
 aleatorios, 43-57
 de tratamiento, 43
 de efectos
 aleatorios, 44, 58, 90, 94
 fijos, 44, 58, 90, 94
 mixtos, 90, 94
 de encuesta, 8, 10, 13
 de grupos completamente aleatorios, 22, 25
 de interacción, 282
 de investigación, 1
 de $N > 1$, 21
 de medida única (*véase* diseño intersujetos)
 de medida múltiple (*véase* diseño de medidas repetidas y diseño de medidas parcialmente repetidas)
 de medidas parcialmente repetidas, 22, 282-302
 de medidas repetidas, 22, 25, 42, 241-320
 de muestra (*véase* diseño *split-plot*)
 de parcela dividida (*véase* diseño *split-plot*)
 de perfiles, 283
 de replicación fraccionada (*véase* diseño factorial fraccionado)
 de sujetos intragrupos intratratamientos (*véase* diseños jerárquicos)
 de tratamientos $\times$ sujetos (*véase* diseño de medidas repetidas)
 emparejados, 23, 25, 157
 equilibrados, 24, 26
 experimental, 1-16
 clásico, 21, 27
 clasificación del, 21-26
 factorial, 22, 25
 aleatorio, 89-136

clasificación, 89-90
 completo, 90
 de base fija, 89
 de medidas repetidas, 241-243
 fraccionado, 180, 346-355
 incompleto, 90
 mixto, 282-302
 fisheriano, 21, 27
 fraccionados, 23, 26, 42
 funcionales, 22, 25, 57
 incompletos, 23, 26, 190
 accidentalmente, 23, 26
 estructuralmente, 23, 26, 42
 intergrupos o intersujetos, 22, 25, 42
 intrasujeto (*véase* diseño de medidas repetidas)
 jerárquicos, 23, 25, 26, 190-208
 Lindquist tipo I, 282
 longitudinal, 8
 mixto (*véase* diseño de medidas parcialmente repetidas)
 multigrupos, 22, 25
 aleatorios, 57-66
 multinivel, 22, 25
 multivalentes, 22, 25
 multivariado, 23, 25, 42, 321-332
 multivariante (*véase* diseño multivariado)
 no equilibrados, 24, 26, 75
 no-experimental, 8, 11, 43
 no paramétricos, 24, 26
 observacional, 8, 10-11
 paramétricos, 24, 26
 parcialmente aleatorizados (*véase* diseños de bloques aleatorios)
 simple (*véase* diseño unifactorial)
 de medidas repetidas, 241-243
 sin covariables, 24, 26
split-plot, 282-302
 transversal, 8
 unifactorial, 22, 25
 univariado, 23, 25
 univariante (*véase* diseño univariado)

DMS de Fisher (*véase* Fisher)

Duncan (*véase también* comparaciones múltiples)
 método de, 68, 70, 74

Dunn (*véase también* comparaciones múltiples y Bonferroni)
 prueba de, 68, 71

Dunn-Sidák (*véase también* comparaciones múltiples)
 prueba de, 71

Dunnet (*véase también* comparaciones múltiples)

procedimiento de, 34, 70, 73, 81-82, 84
 prueba C de, 76
 prueba T3 de, 76

E

Efectos

carry-over, 182, 186, 242, 309, 314
 de interacción, 90-91, 113, 121, 135-136, 142-151, 334, 337
 de período, 242
 factoriales, 24, 333, 337
 principales, 90, 113, 334, 337
 residuales (*véase* efectos *carry-over*)
 simples, 91, 142-151, 334

Eigenvalue (*véase* autovalores)

Emparejamiento

técnica de, 23, 157

Equilibración

técnica de, 23, 42

Error

de tipo I, 35, 46, 91, 150, 212, 243-244, 246
 de tipo II, 46
 de tipo IV, 91
 experimental, 28, 159

Experimental

diseño (*véase* diseño experimental)
 investigación (*véase* diseño experimental)
 metodología (*véase* metodología experimental)
 Psicología, 7

Experimento, 14, 16

de campo, 7
 de laboratorio, 7

F

F conservadora, 46, 245, 283

de Newman-Keuls (*véase* Newman-Keuls)
 de Ryan, Einot, Gabriel y Welsch (REGWF)
 (*véase también* comparaciones múltiples), 72

F* de Brown y Forsythe, 46

Factor (*véase* variable independiente)
 anidado (*véase* variable anidada)
 de anidamiento (*véase* variable de anidamiento)

Fisher (*véase también* comparaciones múltiples)
 diferencia menor significativa de (LSD o DMS), 73
 método de, 68, 70

Fisher-Hayter (*véase también* comparaciones múltiples)
prueba de, 74
Función discriminante, 37

G

Games-Howell (*véase también* comparaciones múltiples)
prueba de, 76
Grados de libertad, 32, 37, 52
de error, 32, 53, 182, 186-187
intergrupos, 32, 52
totales, 32, 53
Greenhouse y Geisser ($\hat{\epsilon}$), 244-245, 278, 283, 302, 312

H

Hartley
prueba de, 46-47, 60
Hipótesis, 16
alternativa, 11, 17, 36, 113, 213-214
causal, 13
de covariación, 13
de nulidad (*véase* hipótesis nula)
estadística, 17
explicativa, 8
rival, 9
inferencia de la, 8-9
nula, 17, 33, 36, 214
operacionalización de, 5
Hochberg GT2 (*véase también* comparaciones múltiples)
prueba de, 75
Holm-Shaffer (*véase también* comparaciones múltiples)
prueba de, 74
HSD de Tukey (*véase* Tukey)

I

Ideográfico, 7
Interacción (*véase* efectos de interacción)
Investigación
aplicada, 17
experimental (*véase* experimental)

J

Ji-cuadrado, 44

K

Kolmogorov-Smirnov
prueba de, 46, 54

L

Lambda de Wilks, 38, 40, 325-329
Levene
prueba de, 46, 55, 87, 109, 140, 207
Lilliefors
prueba de, 46
LSD de Fisher (*véase* Fisher)

M

Manipulación, 13-17
MANOVA (*véase* análisis de la varianza multivariado)
MAX-MIN-CON, 18
Media cuadrática, 32
Método
cualitativo, 7
cuantitativo, 7
de los intervalos de confianza simultáneos de Roy, 42
de los mínimos cuadrados, 94
Metodología
de encuesta (*véase* metodología selectiva)
experimental, 5-6, 13
naturalista (*véase* metodología observacional)
observacional, 5-6
selectiva, 5-6
Miller-Winer (*véase también* comparaciones múltiples)
prueba de, 75
Mínimos cuadrados ordinarios, 28-29, 94
Modelo
aditivo, 92, 159, 184, 246
mixto univariante, 244
multivariante, 244, 246
no aditivo, 92, 159, 250

N

Newman-Keuls (*véase también* comparaciones múltiples)
método de, 68, 70, 72
F de (*véase* Newman-Keuls método de)
Nivel de significación, 46
Nomotético, 7

O**O'Brien**

procedimiento de, 46

Outliers (véase valores extremos)**P****Paradigma**, 13

asociativo, 13

experimental, 8, 13

Parámetro aproximado de número de medias (véase también Cicchetti), 154**Peritz** (véase también comparaciones múltiples)

prueba de, 70, 74

Polinomios ortogonales, 58**Potencia**, 35, 210

del diseño, 19, 157

estadística, 39, 92, 141, 154, 212, 242, 245, 280

estimación de, 55, 107, 138, 188, 205, 238, 330, 342, 353

para cualquier par de comparaciones, 70

para todos los pares de comparaciones, 70

para un par de comparaciones, 70

Prueba

de Box, 244

de esfericidad de Mauchley, 244, 246, 258, 277

de homogeneidad de varianzas, 46, 55, 85, 107, 138, 205

de la hipótesis, 92

de no aditividad (véase Tukey)

de significación general (véase razón *F*), 67

de tendencias de Jonckheere, 58

LSD, 41

R**Raíz mayor de Roy**, 38, 39, 40, 329**Rango crítico entre pares de medias**, 80-84, 152, 280**Razón *F***, 33, 37, 67**Regla de asignación**

aleatoria, 8, 14-15, 22

Regresión múltiple, 92-93**Robustez**, 39**Roy-Bose**

procedimiento de, 281

S**Scheffé** (véase también comparaciones múltiples)

prueba de, 34, 68, 72, 84, 155, 281

Sensibilidad

de la investigación, 19

Shaffer-Ryan (véase también comparaciones múltiples), 73**Shapiro-Wilk**

test de, 46, 54

Simetría compuesta o combinada (véase también supuesto de esfericidad), 244**Suma**

cuadrática (véase también suma cuadrática)

de error (véase suma cuadrática residual)

intragrupo, 31, 36-37

intergrupos, 31, 36-37

residual, 36

total, 31, 36

de cuadrados, 30

de productos cruzados, 36

Supuesto

de circularidad (véase supuesto de esfericidad)

de esfericidad, 243-246, 279-280, 283

de fiabilidad en la medida de la covariable, 212, 222

de homocedasticidad (véase supuesto de homogeneidad de las varianzas)

de homogeneidad (véase también prueba de homogeneidad de varianzas)

de las matrices de covarianza, 40, 247

de las pendientes de regresión, 212, 220-222, 235

de las pendientes intragrupo (véase supuesto de homogeneidad de las pendientes de regresión)

de las varianzas, 45, 46, 55, 142, 210, 283

de independencia

de las observaciones (véase supuesto de independencia de los errores)

de los errores, 45, 211, 242, 283

entre la variable independiente y la covariable, 211, 219, 235

de linealidad entre la variable dependiente y la covariable, 211, 219

de normalidad (véase también prueba de normalidad), 45, 54, 210, 283

multivariable, 40

de medida de la variable dependiente, 45

de selección aleatoria de los niveles de tratamiento, 45

T**t de Student**, 44, 47, 71

para muestras relacionadas, 280

T² de Hotelling, 42**T³ de Dunnett** (*véase* Dunnett)**Tamaño del efecto**, 35, 107, 138, 188, 205, 210, 238, 242, 330, 342, 353**Tasa de error**

por comparación, 68

por familia, 68

referida a la familia o al experimento, 68

Técnica

de confusión, 333

completa, 333-336

parcial, 336-338

de la *F* o de la *t* protegida, 41

de replicación fraccionada, 347-350

Traza

de Hotelling-Lawley, 38-40, 329

de Pillai-Bartlett, 38-40, 329

Tukey (*véase también* comparaciones múltiples)

prueba de no aditividad, 163, 314

prueba DHS (o HSD), 34, 68, 70, 71, 82-84, 153-155

prueba WSD, 280

Tukey-Kramer (*véase también* comparaciones múltiples)

prueba de, 75

U**U de Mann-Whitney**, 44**V****Validez**

externa, 43, 91-92

interna, 8, 43

de conclusión estadística, 91, 242

de la investigación, 19

Valores extremos, 75-76**Variable**

anidada, 190-197

atributiva, 282

de anidamiento, 190-197

de bloqueo (*véase también* bloqueo), 158de confundido (*véase también* variable extraña), 20

de efectos aleatorios, 44

de efectos fijos, 44

de emparejamiento, 157

dependiente, 14-15, 18

extraña, 16, 18-20, 158, 190, 208, 282

independiente, 14-15, 18, 22

perturbadora (*véase* variable extraña)

pronóstica, 282

Variación (*véase también* varianza)

explicada, 36

no explicada, 36

Varianza

asistématica, 19

de error, 18-19, 44, 92, 158, 242

experimental, 18

intergrupos, 18

intragrupo, 19

no pretendida, 18

pretendida, 18

residual (*véase* varianza de error)

sistemática, 18

primaria, 18-19

secundaria, 18-20, 191

total, 18

W**W de Welch**, 46**Y****Yates, notación de**, 333-334

Diseños de Investigación Experimental en Psicología

Balluerka • Vergara

Diseños de Investigación Experimental en Psicología es un libro dirigido a estudiantes y profesionales de Psicología y de otras Ciencias del Comportamiento. Su objetivo radica en proporcionar al lector las pautas necesarias para realizar investigaciones originales utilizando diseños experimentales.

El libro comienza abordando el concepto de diseño, es decir, describiendo los elementos que forman parte de la planificación de la investigación desde un punto de vista formal. Tras esta aproximación general al diseño, se exponen los principales criterios de clasificación de los múltiples diseños experimentales que se utilizan actualmente en las Ciencias Sociales y del Comportamiento. En la parte más extensa de la obra se abordan de forma muy didáctica y mediante numerosos ejemplos cada modelo de diseño y la técnica de análisis estadístico asociada al mismo. El desarrollo matemático de los análisis se complementa con la presentación del procedimiento necesario para llevarlos a cabo mediante el paquete estadístico SPSS 10.0.

En definitiva, se trata de proporcionar al lector un manual de diseños experimentales que no constituya únicamente un referente teórico sino, sobre todo, un referente aplicado que le permita optimizar la calidad de sus investigaciones.

Nekane Balluerka y Ana Isabel Vergara son profesoras de Diseños de Investigación en la Facultad de Psicología de la Universidad del País Vasco.

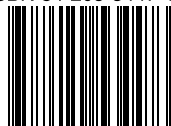
LibroSite es una página web asociada al libro, con una gran variedad de recursos y material adicional tanto para los profesores como para estudiantes. Apoyos a la docencia, ejercicios de autocontrol, enlaces relacionados, material de investigación, etc., hacen de LibroSite el complemento académico perfecto para este libro.

www.librosite.net/balluerka



www.pearsoneducacion.com

ISBN 84-205-3447-1



9 788420 534473

ISBN13: 978-84-205-3447-3